Regulation Of Transcription In Maedi-Visna Virus.

Keith A. Sutton

This thesis is submitted as part of the course requirements for the degree of Doctor of Philosophy at the University of Edinburgh.

January, 1995.

ACKNOWLEDGEMENTS

I would like to express my gratitude to everyone who has offered advice during the course of this work and in particular my supervisors Gordon Harkiss and David Sargan. I am also grateful for the help and advice provided by John Quinn, Doug Roy, Carolyn Fiskerstrand, Ian Bennet and Lawrence Dickson. I would also like to thank my friends and family for their support and encouragement over the last three and a half years.

I am grateful to the MRC for their finacial support which enabled me to carry out this work.

DECLARATION

The compoeition of this thesis and the experiments described are my own work, unless specifically stated in the text. No part of this work has been, or will be, submitted for any other degree, diploma or qualification.

Keith Anderson Sutton

Contents.

Abbreviations

ATP	adenosine triphosphate
APS	ammonium persulphate
ß-Gal	ß-galactisidase
bp	base pairs
BrdU	5-bromo deoxyuridine
CAEV	caprine arthritis encephalitis virus
CAT	chloramphenicol acetyltransferase
Ci,µCi	Curie, microCurie
CIP	calf intestinal phosphatase
cpm	counts per minute
CTL	cytotoxic T lymphocyte
DMSO	dimethylsulphoxide
DNA	deoxyribonucleic acid
DTT	dithiothreitol
EDTA	ethylenediamine tetra-acetate
FIV	feline immunodeficiency virus
HIV	human immunodeficiencu virus
HTLV	human T lymphotrohpic virus
Ig	immunoglobulin
kb	kilobase
kDa	kiloDalton
LB	Luria broth
LTR	long terminal repeat
M,mM,µM	molar, millimolar, micromolar
ml,µl	millilitre, microlitre
mRNA	messenger RNA
MVV	maedi-visna virus
ONPG	2-nitrophenyl-ß-galactopyranoside
PBM	peripheral blood monocytes
PCR	polymerase chain reaction
RNA	ribonucleic acid

SA-OMVV	South African ovine maedi visna virus
SDS	sodiumdodecyl sulphat
SDW	sterile distilled water
TAE	tris-acetate-EDTA
TBE	tris-borate-EDTA
TEMED	tetrametylethylenediamine
UV	ultraviolet

CHAPTER ONE

INTRODUCTION

1.1 Maedi-Visna Virus: an Ovine Lentivirus

Maedi-visna virus (MVV) was initially isolated and characterised following outbreaks of progressive paralysis (visna) and progressive pneumonia (maedi) in Icelandic sheep flocks between 1939 and 1952. Transmission of the disease to healthy animals using tissue preparations from infected animals was demonstrated (Sigurdsson 1954a, Sigurdsson *et al.*, 1957) and in 1967 a single virus was shown to be responsible for both symptoms (Gudnadottir & Ralsson 1967). The long interval between MVV infection and onset of MVV mediated disease was used to define a new form of virally induced disease defined as 'slow' (Sigurdsson 1954b). This was to distinguish the disease process from acute viral infection.

MVV is thus the prototypic member of the Lentivirinae subfamily of the Retroviridae family of viruses. The Lentivirinae subfamily includes, in addition to MVV, caprine arthritis-encephalitis virus (CAEV, Crawford *et al.*, 1980), equine infectious anaemia virus (EIAV, Dreguss & Lombard 1954), feline immunodeficiency virus (FIV, Pedersen *et al.*,1987), bovine immunodeficiency virus (BIV, Gonda *et al.*, 1987), simian immunodeficiency virus (SIV, Letvin *et al.*, 1985) and Human Immunodeficiency Virus types 1 and 2 (HIV-1 and HIV-2). With the exception of EIAV which causes cyclical disease episodes, the lentiviruses cause progressively debilitating disease, usually with slow onset.

Phylogenetic comparisons of the lentiviruses differ in fine detail dependent on the gene and viral strains studied. These studies do, however, provide a reasonable consensus that MVV is more closely related to the non-primate lentiviruses than to the primate species (Olmsted *et al.*, 1986, McClure *et al.*, 1988, Olmsted *et al.*, 1989). Like all members of the retroviridae, MVV possesses a positive sense single stranded RNA genome (ssRNA) (Brahic *et al.*, 1977), an RNA dependent DNA polymerase (Lin & Thormar 1970) and replicates via a DNA intermediate, the provirus (Haase & Varmus 1973). Although possessing a ssRNA genome the retroviruses package two copies of the genome per virion. Thus MVV, like all lentiviruses, is described as being diploid. This packaging of two genomes is one of the features which allows the retroviruses to achieve high mutation rates (reviewed by Coffin 1979, Katz & Skalka 1990). Recombination between the two genomes during reverse transcription allows sequences within 10s of bases to behave as if they were unlinked.

1.2 Genomic Organisation of MVV

The lentiviruses possess a complex genome structure which, in addition to the *gag*, *pol* and *env* genes, contains a number of small open reading frames (orfs). These orfs encode auxillary proteins involved in the control of viral replication.

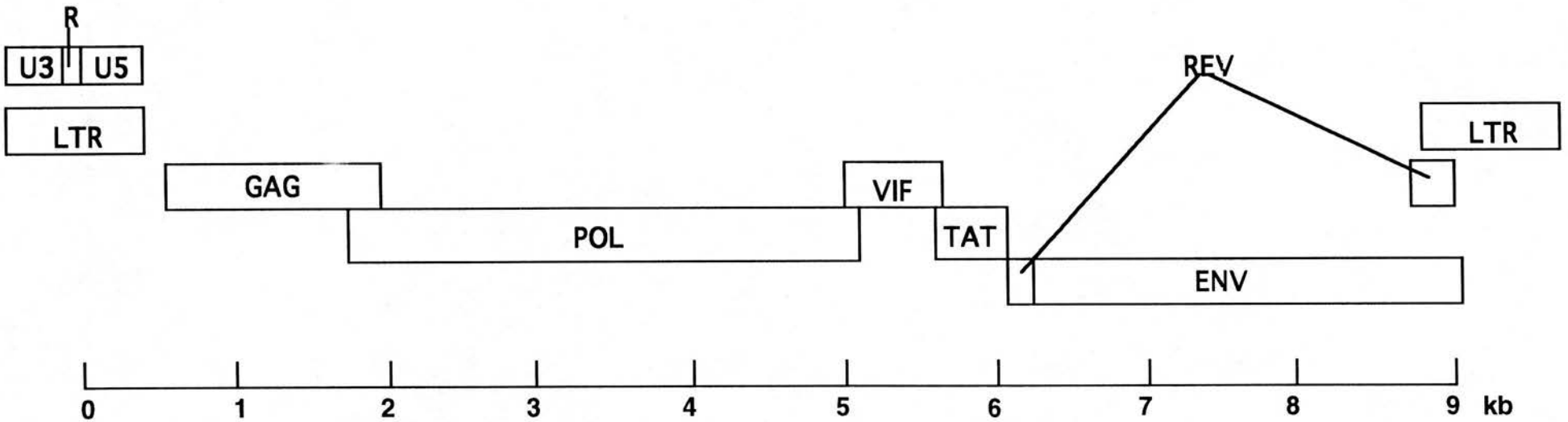
Seven MVV isolates have been sequenced to date. Four of these are variants of the 1514 strain (Sonigo *et al.*, 1985, Braun *et al.*, 1987, Staskus *et al.*, 1991, Andresson *et al.*, 1992) while the others are South African-Ovine Maedi Visna Virus (SA-OMVV, Querat *et al.*, 1990) and the British isolate EV-1 (Sargan *et al.*, 1991). These sequences were all derived from molecular clones with the exception of the EV-1 sequence which is from a biological clone. The viruses of these 3 strains possess a similar genomic structure with only minor differences. A basic outline of the MVV genome is shown in Fig 1.1. The DNA genome of the lentiviruses is flanked by two long terminal repeats (LTRs). The LTR sequence is composed of the U3, R and U5 regions. This sequence contains a primer binding site complementary to that for $tRNA^{lys}_{1,2}$ (Sonigo *et al.*, 1985) and a polypurine tract which acts as the initiation site for positive strand DNA synthesis. The LTR also contains the target sequences which are responsible for regulating transcription of the viral genome (Section 1.6).

The MVV viruses appear to have a less complex pattern of orfs than HIV-1, the most extensively studied lentivirus. In this thesis, the HIV orf nomencluture will be used. Potential homologues to the HIV orf proteins Tat (Gourdou *et al.*, 1989) , Rev (Tiley *et al.*, 1990) and Vif (Strebel *et al.*, 1987, Audoly *et al.*, 1992) have been described in MVV. In addition to these three orfs, which are known to produce functional proteins, SA-OMVV contains a novel orf termed W (Querat *et al.*, 1990). There is, however, no evidence for translation of this sequence. It should be made clear that although some of the MVV orf products are functionally homologous to the HIV proteins (Rev and Vif) others may be acting in a quite distinct manner (Tat). MVV does not appear to encode homologues for the HIV-1 orfs *vpr* (Ogawa *et al.*, 1989), *vpu* (Strebel *et al.*, 1988), *tev* (Benko *et al.*, 1990) or *nef* (Ahmed & Venkatesan 1988). The regulatory orf proteins of MVV will be discussed in greater detail in Section 1.5.

Figure 1.1

An outline of the MVVgenome is illustrated. The locations of the 6 translated open reading frames are indicated. All sequenced isolates of MVV have a size of between 9 and 9.3kB. The LTR is divided into U3, R and U5 regions. The promoter sequences controlling transcription are present within the U3 region.

FIGURE 1.1. - The MVV Genome



1.3 Pathology of MVV infection

During the original outbreak of MVV in Iceland, an incubation period of between one to three years was described, although in some cases the incubation period was longer (Sigurdsson *et al.*, 1954a). This prolonged asymptomatic period was followed by the onset of clinical disease. This was characterised by a gradual progression of symptoms, with little regresssion, until death occurred. A number of organ systems were affected most notably the lung, mammary gland and central nervous system (CNS). The lesions were characterised by their inflammatory nature and lymphocytic infiltrate.

The lung disease causes dyspnoea (Maedi). At post-mortem the lungs may be two to four times their normal weight with decreased elasticity and a degree of fibrosis (Sigurdsson *et al.*, 1954a). Accumulation of plasma cells, mononuclear phagocytes and lymphocytes leads to a thickening of the alveolar septa (Georgsson & Palsson 1971). In severe disease this infiltrate generates lymphoid follicles with active germinal centres (Lairmore *et al.*, 1986). The regional lymph nodes become enlarged.

Neurological infection affects gait with paralysis of the hindquarters progressing until paraplegia develops (Visna). Infection of the CNS leads to demyelination and destruction of the white matter in the brain, cerebellum and spinal chord (Sigurdsson *et al.*, 1957). Primary inflammatory lesions appear to be focused on the glia with little neuronal destruction (Sigurdsson & Palsson 1958, Stowring *et al.* 1985). Infection of oligodendrocytes and astrocytes by MVV has been observed *in vivo* (Stowring *et al.*, 1985). In the mammary gland, MVV generates a chronic indurative mastitis (Van der Molen *et al.*, 1985) which is again characterised by marked lymphoid hyperplasia and slight fibrosis as occurs in the lung. One additional site of pathology in some animals is the joint. Swelling and calcification of soft tissue, fibrosis of the joint capsule and synovium and perivascular lymphocytic infiltration leads to arthritis (Oliver *et al.*, 1981). This pathology is similar to that seen with CAEV (Crawford *et al.*, 1980). In all these sites of pathology only very few cells are seen to be expressing viral proteins. The mechanisms controlling viral pathogenesis are poorly understood and the events which lead to the initiation of the viral lesions and pathology are unknown. The lifecycle of MVV has been analysed with regard to identifying steps which are linked to the observed pathology. This can be simplified to two major questions; firstly, how does MVV evade the immune response during the long incubation period and secondly, how does the virus

dysregulate normal homeostasis and induce these massive cellular infiltrates? There is as yet no firm answer to either of these two questions, but *in vitro* and *in vivo* studies on the viral lifecycle and regulatory genes have begun to provide clues as to possible mechanisms.

1.4 The MVV Lifecycle

In vivo, MVV appears to infect preferentially cells of the monocyte and macrophage lineage (Gendelman *et al.*, 1985, Peluso *et al.*, 1985) although numerous other tissues have been shown to be susceptible to infection *in vivo.* These include choroid plexus cells (Brahic *et al.*, 1981), lung epithelia (Geballe *et al.*, 1985) and both oligodendrocytes and astrocytes in the CNS (Stowring *et al.*, 1985). The nature of the MVV receptor remains to be fully characterised (Crane *et al.*, 1991, Dalziel *et al.*, 1991). Given the range of cells apparently susceptible to infection it appears to be widely distributed.

After binding the cellular receptor the MVV particle fuses with the cell. The virus then enters and the virion is uncoated. Reverse transcription of the viral nucleic acid then proceeds, primed off the cellular $tRNA^{lys}$ bound to the LTR. The viral Pol protein then generates a linear dsDNA copy of the genome which is then circularised. This circularised form is thought to be topologically preferred for integration into the host cell's DNA (Panganiban & Temein 1984).

1.4.1 Integration

The confirmation of integration as a critical step in the MVV lifecycle still awaits a full experimental demonstration. However, integration has been shown in a variety of systems to be a critical step in the lifecycle of the retroviruses (reviewed by Coffin 1990). The importance of integration during *in vitro* MVV infection has been questioned (Harris *et al.*, 1985). These workers found no evidence for integrated genomes during lytic infection of sheep choroid plexus cells. However, using different permissive cells, Vigne *et al.* (1985) observed integration in a minority of cells. This reported failure to observe large numbers of integrated genomes does not appear to be due to a defect in the MVV integrase protein. Studies using purified recombinant MVV integrase have demonstrated that it is competent to perform the full

repertoire of *in vitro* functions characteristic of retroviral integrases (Katzman & Sudol 1994). Recent experiments with integrase deficient molecular clones suggest that these viruses are non-infectious (G.Querat personal communication). Thus, while during lytic *in vitro* infection there may only be a small number of integrated genomes these may be vital for the viral lifecycle.

In HIV, where integrase deficient viruses have been functionally tested, there is some disagreement as to the consequences of integrase deletion. Several studies have suggested that the integrase activity is an absolute requirement for productive infection (Stevenson *et al.*, 1990a, LaFemina *et al.*, 1992, Sakai *et al.*, 1993). These studies do, however, disagree as to the mechanism by which the integrase mutation blocks the infectious process. The data of Stevenson *et al.* appears to suggest a late action of the defect, as expression of viral proteins was unimpaired. This would suggest a novel function for the Integrase as mutants appear to generate a phenotype similar to Vif mutants (Section 1.5.2); perhaps affecting virion assembly or some other late process. In contrast, the data of Sakai *et al.* is consistent with observations in other retroviruses such as murine leukemia virus (Donehower & Varmus 1984, Roth *et al.*, 1990) and Rous Sarcoma Virus (Hippenmeyer & Grandgenett 1984) where mutation of the integrase blocked viral gene expression. One problem of studies on the integrase protein is that it is produced as a precursor polyprotein which is cleaved to generate the mature Pol and Gag proteins. As the Pol protein contains a number of distinct enzymatic activities any mutation within the integrase domain may interfere with the activity of these other sites (Roth *et al.*, 1990, Stevenson *et al.*, 1990). Using the HIV-1 integrase Shin *et al.* (1994) identified a number of single point mutations within the integrase active site which failed to block integration but which did affect virion assembly and infectivity. Additional studies by this group identified single amino acid substitutions which blocked the production of virus but these viruses were not integration defective (Taddeo *et al.*, 1994). Further it was possible to identify mutants which blocked integration but did not block replication. Using a similar single amino acid substitution approach Wiskerchen & Muesing (1995) failed to observe replication of integrase mutants. These workers tested the producer cell lines for proper expression and processing of the Pol and Gag proteins. From this work it was concluded that integration was essential for viral replication in primary cell lines.

The integrase activity of the Pol gene has been accepted as playing a central role in the retroviral lifecycle. The studies described above suggest that mutations within the

integrase domain can block the activity of other domains within the Pol protein so blocking replication. This may account for the some of discrepencies described in previous studies on the phenoytype generated by the deletion of integrase activity in HIV-1 (Stevenson *et al.*, 1990, Sakai *et al.*, 1993).

The reason for the dependence of viral gene expression on integration is unclear. On balance it is accepted that retroviral replication is dependent on integration, and linearization, within the host genome. This dependency on integration for gene expression is not due solely to the topology of the genome as circularized DNA is transcribed as efficiently, if not more so, than the linear form (Harland *et al.*, 1983). Thus the failure of unintegrated retrovirus genomes to replicate would appear to be a property of the genome, or the nucleoprotein complex which forms on it. The role of viral integration in the lifecycle is also of importance when it comes to analysing LTR function. Jeang *et al.* (1993) have reported that integrated HIV LTRs show a distinct activity and Tat responsiveness when compared to unintegrated genomes. One interesting result of these studies was that transient co-transfection assays had suggested that exon 1 of Tat was as active as proteins containing both exons. Jeang *et al.* have demonstrated that this effect was due to the non-integrated nature of the target vectors. When integrated LTRs are studied the two exon form of Tat was up to 10 fold more active than the single (1st) exon protein. This highlights possible complications, and the danger of over interpretation, of model systems.

The provirus possesses an LTR, and hence promoter sequences, at both ends of the genome. In the integrated HIV genome the 5' LTR directs initiation of transcription and sequences in the 3' LTR control cleavage and polyadenylation. The basis for this promoter selection is unclear as both LTRs share a common nucleotide sequence (Ju & Culllen 1985). Studies on this asymmetry in HIV have shown that both LTRs are transcriptionally active and Tat inducible. However, the 5' LTR drives a much higher level of transcription than the 3' LTR (Klaver & Berkhout 1994). These studies showed that both LTRs are functioning suboptimally. Transcription from the 5' LTR can be elevated by deletion of the 3' LTR and likewise deletion of the 5' LTR activates the 3' LTR. The basis for this promoter suppression is unknown.

It appears that while HIV can infect and persist in unactivated T cells, activation of the cells is required for integration and the triggering of efficient gene transcription (Stevenson *et al.*, 1990b). These resting cells were able to maintain extrachromosomal HIV genomes for 2 weeks *in vitro.* The maximum period that these unintegrated genomes can remain in resting

cells has not been determined. Integration may thus represent a potential block in the normal HIV lifecycle *in vivo*. Rather than integrating and becoming latent, the virus remains in a unintegrated state, which may be blocked for transcription and replication, until cell activation. This provides a novel mechanism to prevent premature gene expression which may reveal the infected cell to immune surviellance. The relevance of these observations to MVV has yet to be determined but they do provide a possible insight as to the sequence of early events in lentiviral infection and their possible significance to the processess of viral replication and latency.

Studies on cell lines non-productively infected with HIV have led to the proposition that integration site selection may affect viral gene transcription and play a role in latency (Winslow *et al.*, 1993). These non-productively infected cell lines appear blocked at the early non-productive stage of infection and multiply spliced RNAs (encoding Tat, Rev and Nef) predominate over singly or unspliced RNAs (Pomerantz *et al.*, 1990). These cell lines can be induced to produce virus following activation with phorbol esters or lipopolysaccharide. The endogenous virus produced from these cells is infectious showing that the lack of efficient replication is not due to a defective genome. The cell lines which show this blocking of endogenous virus can support the replication of exogenous virus under the same conditions where production of virions by the endogenous virus is blocked (Winslow *et al.*, 1994). This superinfection with a second HIV virus leads to production of virus derived from both the exogenous and endogenous viruses. This activation of the endogenous virus is believed to occur in *trans* via Rev produced from the exogenous virus. Winslow *et al.* concluded from these observations that the integration site of the endogenous virus was preventing efficient transcription and blocking the accumulation of Rev and other products of the multiply spliced RNAs which control the switch to virus production. Thus the location of the integration site in the host genome may affect the outcome of viral infection for a given cell.

The integration site within the host genome is selected in a non-random manner, with structure of the host genome affecting integration apparently more than the presence of an integration target sequence (reviewed by Craigie 1992). Retroviral integration does not appear to be sequence specific *in vitro* or *in vivo*. However, *in vivo*, integration of retroviruses appears to be preferentially directed to sites where the chromatin is sensitive to nuclease digestion (Shih *et al.*, 1988). There is also other evidence that retroviruses integrate preferentially into transcriptionally active regions of the host genome. Hwang & Gilboa (1985) compared the

efficiency of transfection and retroviruses for the transfer, and expression, of a foreign gene in tissue culture cells. The retroviral transfer gave levels of transcripts 10-50 fold higher than the levels acheived with calcium phosphate transfection. These workers concluded that their results were consistent with an elevated integration of the retroviral vector, when compared to the transfected vector, into transcribed regions.

There is evidence that DNA bending results in the generation of favoured sites for retrovirus integration (Pruss *et al.*, 1994, Muller & Varmus 1994). One consequence of nucleosome formation is DNA bending thus creating optimal sites for retrovirus integration. It is important to note that this insertion is triggered by DNA bending itself and not the DNA/nucleosome interaction (Muller & Varmus 1994). This and other and other evidence (reviewed in Sandmeyer *et al.*, 1990) suggests preferential integration into transcribed regions of the host genome. This makes sense as the integration of the virus in regions where transcription is inactive is likely to lead to silencing of the genome, due to chromatin condensation, and failure of the virus to propagate itself.

These early steps in the retroviral lifecycle i.e. the generation of the DNA genome and integration, are critical for later steps of the lifecycle. The failure to integrate or integration at a heterochromatic locus will block all further stages of the lifecycle. However, it does appear that lentiviruses may be able to take advantage of the transcriptionally silent state of unintegrated genomes and use this as a mechanism for generating latent infection.

1.4.2 Viral Replication

After these early events in the infection process, reverse transcription generating the dsDNA genome and integration, the process of viral replication follows distinct patterns *in vivo* and *in vitro*. *In vitro* MVV infection of permissive cells, such as sheep choroid plexus, gives rapid viral replication with a pronounced cytopathic effect (CPE) and the release of free infectious virus into the culture medium. *In vivo* there appears to be a block to rapid replication and this state has been referred to as restricted replication (Haase *et al.*, 1977, Brahic *et al.*, 1981). The nature of this *in vivo* block of viral replication is poorly understood and it is unclear what the relative contribution of viral and cellular products is in this process. Some form of restricted or latent state is likely, however, to play a key role in the persistance of infection within the animal.

1.4.3 Restricted Replication

Through this thesis the concept of 'restricted replication' will be referred to on a number of occasions. This concept is based on the idea that the viral genome can direct expression of only the small spliced transcripts (Tat and Rev in MVV) without the expression of structural genes. Restricted replication is thus proposed as a possible mechanism of lentivirus persistance where viral protein expression is kept below levels which the immune system can detect (Haase *et al.*, 1977, Haase *et al.*, 1994). The ideas of 'restricted replication' and 'latency' are therefore distinct. However, if a state of restricted replication does exist, it is possible that the viral genome may move from a latent state to a restricted state and then either back to latency or on to a fully activated state. The actual definition of restricted replication is unclear, as the term is often used to cover any situation *in vivo* where the infected cell is expressing the viral genome at levels below that seen in permissive infection *in vitro* (Geballe *et al.*, 1985). The most meaningful definition is probably that only those cells expressing the small spliced transcripts (i.e. Tat & Rev) and no structural proteins are in a state of restricted replication. Any cell producing detectable quantities of structural proteins can no longer be said to possess the viral genome in a truly restricted state and is likely to be producing infectious virus. Such cells are more likely to be subject to immune surveillance due to the expression of the viral structural proteins as these appear to be the prime targets of the both the antibody and cytotoxic T cell response in MVV and other lentiviral infection (Clerici *et al.*, 1989, Schrier *et al.*, 1989, Knight *et al.*, 1992, Nixon 1992, Blacklaws *et al.*, 1994). It is probable that there will be a population of T cells recognising the antigens derived from the small multiply spliced RNAs in MVV infected animals but the prevelance of these cells remains to be determined. In MVV there is some indirect evidence to support the idea of a small population of infected cells being in a state of restricted replication; that is viral gene expression with no expression of structural proteins (Brahic *et al.*, 1981). However, when more sensitive PCR techniques are used then it becomes clear that the major population of virally infected cells, in both MVV and HIV, have viral genomes present in a truly latent state (Embretson *et al.*, 1993, Haase *et al.*, 1994). One potential problem of the earlier combined *in situ* /immunocytochemistry studies is that the sensitivity difference between the two techniques meant that many infected cells producing viral RNA may in fact have been expressing viral structural protein.

Studies using lymphoid tissue from HIV infected individuals found that message for

the *gag* structural gene was expressed throughout infection (Pantaleo *et al.*, 1993). The conclusion of these workers was that, although replication of the virus was site dependent, with the lymphoid tissues showing higher replication than blood at all stages except late in disease, throughout the latent period HIV infection is active and productive. This study did not claim the absence of a restricted or latent state but rather during that period of infection termed 'latent' when there is no discernable disease lentiviral replication is proceeding at a steady and detectable rate.

These various studies suggest that the majority of infected cells are not in a state of restricted replication but rather, are tru ly latent with no detectable gene expression. There may be a pool of cells in a 'restricted state' (Embretson *et al.*, 1993). However, it is likely that this represents a short transition phase between latency (or initial infection as there may be no requirement for a latent phase in the viral lifecycle) and production of viral particles. Thus it may be more accurate to refer to the situation where low level expression is observed as 'leaky' or 'incomplete' latency in a subpopulation of infected cells rather than as a distinct restricted state. To be consistent with the literature the term 'restricted replication' will be used in this thesis to refer to the state of vi replication where there is gene expression but at levels below those seen *in vitro; ie* detectable viral RNA transcripts but no directly detectable viral protein.

The possible trigger for initiating the transition from latency to active gene expression is unknown. It is critical to elucidate the exact regulation of lentiviral gene expression during infection as this will have an impact on possible theraputic stratagies. If the restricted state is prevalent then vaccination against the products of the early spliced transcripts may allow elimination of cells prior to the generation of infectious virus. On the other hand if there is a rapid transition from a latent state, which cannot be controlled by the immune system, to active replication and the release of infectious virus then it is possible that vaccination strategies will not protect against infection. This second model sits well with the 'Trojan horse' mechanism for lentivirus dissemination through the host (Peluso *et al.*, 1985). Virus produced from an infected cell infects surrounding cells which can then spread through the body, or be transmitted to another host, while carrying the latent viral genome.

1.4.4 Viral Transcription

While little is known about the lentiviral replicative process *in vivo* it has been extensively studied *in vitro*. The dsDNA genome serves as the template for transcription of viral genes and for the full length viral genomic RNA which will be packaged into virions at the end of the lytic cycle. *In vitro* MVV infection leads to the formation of syncytia and cell lysis. Viral transcripts are first detected six to eight hours post infection (Davis *et al.*, 1987, Vigne *et al.*, 1987, Sargan & Bennet 1989). These three studies examining the temporal expression of viral transcripts differ on the exact timing of the appearence of the various RNA species. They do all, however, demonstrate a temporal regulation of spliced and unspliced viral RNAs. The RNAs which accumulate up to 24 hours represent small multiply spliced transcripts which contain sequences from both the 5' and 3' regions of the genome. After the accumulation of these small RNA species, larger singly spliced and unspliced transcripts are generated. These larger forms code for the structural Gag, Pol and Env proteins. The fully spliced, early transcripts, are believed to encode the Tat (Gourdou *et al.*, 1989) and Rev (Tiley *et al.*, 1990) proteins. These regulatory proteins are then capable of modulating viral and cellular gene expression (Tat) and splicing of viral RNAs (Rev).

The lytic, *in vitro*, infection leads to a rapid accumulation of viral genomes within the cell. Up to 4000 copies of the RNA genome have been reported (Brahic *et al.*, 1977). At the end of the lytic cycle the viral genomic RNA, polymerase and structural proteins are assembled into virions. The particles form by budding from either the cell membrane or, in the case of macrophages, into cytoplasmic vacuoles (Georgsson *et al.*, 1991).

A number of distinctions should be made between the viral replicative process *in vitro* and *in vivo*. While virus replication is rapid *in vitro*, it is not usually possible to recover virus directly from the tissues of infected animals. Infected cells can be identified either by growing tissue explants in culture or by culturing the tissues with a permissive cell line. The requirement for *in vitro* cultivation, which leads to cell activation, for the generation of virus is consisitent with a block on viral replication *in vivo*. Neither of these cultivation techniques is totally reliable, which highlights the rarity of permissively infected cells *in vivo*.

Studies to determine the degree of viral replication *in vivo* do differ in the number of infected cells and proportion of these cells which express viral proteins. Examination of choroid plexus cells of experimentally infected lambs (Haase *et al.*, 1977) using *in situ*

hybridisation detected viral DNA in 18% of cells. Viral protein, as measured by p30 staining, was only detectable in 0.1% of the cells containing viral DNA. Consequent studies (Brahic *et al.*, 1981) reported only 1 to 3 % of choroid plexus cells to be infected. These infected cells contained 60-70 copies of viral DNA , with viral RNA present at levels two orders of magnitude below those found *in vitro*. Contrasting with these studies, experiments examining the degree of alveolar macrophage infection in experimentally infected lambs (Gendelman *et al.*, 1985) showed 15% of these cells to contain greater than 1000 copies of viral RNA. These authors estimated that only 1% of these cells were producing virus and so proposed a post transcriptional block on viral replication. No experiments were performed to immunostain the cells used for the *in situ* analysis to look at the expression of viral proteins. It is thus unclear what relationship there was between the quantity of RNA expressed (which varied from 50 to over 1000 copies per cell) and viral protein synthesis.

In the two studies referred to in the preceding paragraph a massive forced infection was used; either to the brain or to the lung. Following this infection these workers studied only the early stages of infection. This will lead to a number of discrepencies between the observations made in these systems and what is seen in natural infection. What is striking is that even with a forcing of the system viral replication appears to be limited and not all available cells are productively infected.

In vivo restricted replication can be observed in monocytes taken from a variety of anatomical locations. Monocytes obtained from blood, lateral ventricles (Peluso *et al.*, 1985) and bone marrow (Gendelman *et al.*, 1985) show viral RNA expression. The observation of infected cells within the bone marrow highlights a possible mechanism for viral persistance and dissemination throughout the body without the need to generate a viremia. The bone marrow cells which harboured viral RNA were not phenotyped so it is not known whether they were mature, resident, monocyte/macrophages or monocyte precursors. Infection of precursors would lead to all progeny monocytes carrying the MVV genome. Even if the infected cells were not precursors but tissue macrophages then these cells would be in a position to infect newly matured monocytes prior to their migration from the bone marrow and seeding to the periphery. One explanation for the variation in detection of cells containing viral RNA and proteins is the dependency of viral gene expression on monocyte maturation. This has been shown *in vitro* where monocyte to macrophage differentiation leads to increased viral gene expression (Gendelman *et al.*, 1986). *In vivo* studies, using transgenes under the control of

the MVV LTR, show that while expression is not limited to the monocyte/macrophage lineage in these cells expression is dependent on differentiation into macrophages (Small *et al.*, 1989 Clements *et al.*, 1994). Thus alveolar macrophages, which are fully differentiated and are likely to be activated, may be capable of supporting greater levels of viral replication than monocytes and macrophages in other tissues. Monocytes and macrophages resident in other sites are likely to be in a lower activation state, unless activated by changes in the cellular environment, e.g. infection with a second virus or cellular activation by cytokines. This dependence on host cell factors may explain the discrepancy between cells containing viral DNA and those producing infectious virus.

1.5 Viral Regulatory Proteins

Three viral regulatory orfs are known to be expressed by MVV (Davis & Clements 1989, Audoly *et al.*, 1992). Using the nomenclature for the HIV orfs these are known as Rev, Vif and Tat. These proteins can be described as functional analogues of the HIV proteins bearing the same names. The Rev and Vif proteins appear to function via similar pathways in both HIV and MVV but the MVV Tat protein appears to functions by a distinct mechanism in the two viruses.

1.5.1 Rev

The MVV Rev protein is translated from two exons within the *env* gene (Fig.1.1). The first exon is in the same reading frame as that which encodes the Env protein. This gives the two proteins a common 48 amino acids at the amino terminus. The second exon of Rev is out of frame with the *env* gene and encodes a unique 119 amino acids (Davis & Clements 1989). Within the various MVV isolates sequenced to date the Rev protein is the least well conserved with between 30 and 35% divergence between isolates (Sargan *et al.*, 1991). As in HIV, deletion of the Rev gene products gives rise to replication incompetent virus (Toohey & Haase 1994). The Rev proteins of HIV and MVV show little sequence homology. The only areas where there is any conservation of amino acids are the basic and leucine rich domains of HIV Rev (Tiley *et al.*, 1991). These domains have, in HIV, been been shown to mediate RNA binding and interaction with cellular factors respectively (Malim *et al.*, 1989, Hope *et al.*, 1990,

Malim *et al.*, 1991). There is evidence to suggest that these two domains are functioning in the same way in the MVV protein (Tiley *et al.*, 1991). *In vitro* analysis of MVV Rev function shows that like the HIV protein it interacts with an RNA sequence present within the *env* gene. This sequence is capable of forming a complex stem-loop structure (Tiley & Cullen 1992). Both the Rev protein and the Rev Responsive Element (RRE) are required for the generation of unspliced and singly spliced transcripts, and the expression of structural genes (Tiley *et al.*, 1990, Tiley & Cullen 1992).

While both HIV and MVV depend upon a Rev/RRE interaction to express structural genes there is no cross recognition between the two systems (Tiley *et al.*, 1990). This is in contrast to what is observed with the HIV Rev and HTLV (Human T-Lymphotropic Virus) Rex proteins. Like the lentiviruses, HTLV-I encodes a protein, Rex, which controls the splicing of full length RNA and export of unspliced RNA from the nucleus (Hidaka *et al.*, 1988). The Rex protein is capable of replacing HIV-1 Rev in *trans* so rescueing Rev mutant viruses and allowing the production of virions (Rimsky *et al.*, 1988). While the MVV Rev protein will not act on the HIV RRE if an MVV RRE is introduced into a Rev defective HIV genome then the expression of MVV Rev will rescue the chimeric HIV virus. There is, however, complete or partial functional reciprocity between the Rev proteins of the MVV strains EV-1 and 1514 and the Rev protein of CAEV (Cork strain)(M. Fotheringham personal communication).

1.5.2 Vif

The *vif* orf, previously described as Q, encodes a 29kD protein of 230 amino acids. This protein is expressed late in the lytic life cycle in tissue culture cells (Audoly *et al.*, 1992). The MVV Vif protein shows 80% homology between isolates (Querat *et al.*, 1990, Sargan *et al.*, 1991). The HIV and MVV Vif proteins show no sequence homology but as in HIV deletion of Vif results in the production of virions with a reduced infectivity (G. Querat personal communication). HIV Vif has been described as a regulator of virion infectivity (Fisher *et al.*, 1987, Strebel *et al.*, 1987). Mutation of *vif* gives rise to virions which are up to 1000 fold less infectious than wild type virus. As there is no detectable Vif protein within the mature virion it appears that it is involved in the assembly of virus; mutation of *vif* appears to interfere with virus assembly. These *vif* defective viruses can be complemented by the expression of *vif* in the producer but not the target cell (Schwedler *et al.*, 1993). These *vif* mutant viruses appeared

defective in their ability to synthesise proviral DNA, even though they contained normal amounts of reverse transcriptase and integrase activities. The MVV protein is similarily distributed to HIV Vif; no detectable protein within the mature virus released from infected cells and a cytosolic localisation within infected cells. The MVV Vif protein is the least well understood of the three orfs known to expressed in MVV.

1.5.3 Tat

Like Rev, the MVV Tat protein is expressed early in infection from a small multiply spliced RNA. Unlike the Rev and Vif proteins there is experimental evidence to suggest that the mechanism of transactivation is distinct to that seen with the HIV Tat protein. This is borne out by evidence which suggests that for *in vitro* growth MVV does not require the Tat protein. (A. Harmache personal communication).

In HIV the Tat protein acts primarily via the TAR (Tat Activated Region) which is present within the R region of the LTR (Dayton *et al.*, 1986, Fisher *et al.*, 1986). The TAR sequence forms a stable stem-loop structure following its transcription. This Tat/TAR interaction is an absolute requirement HIV replication. The TAR region forms a hairpin structure which binds Tat and a number of cellular factors (reviewed by Antoni *et al.*, 1994). In MVV there is no evidence of a sequence within the R/U5 region capable of forming an RNA hairpin structure (TAR) in any of the isolates sequenced to date. Deletion analysis of the R region of the LTR (Hess *et al.*, 1989) identified no sequences in this region which were Tat responsive. Further the Tat proteins of HIV and MVV show no extended homology and have distinct domain structures (Carruth *et al.*, 1994). The absence of a TAR region is not unique to MVV. Both FIV and CAEV show no evidence of TAR regions (Olmstead *et al.*, 1989b, Saltarelli *et al.*, 1990). In the absence of a Tat/TAR interaction the mode of MVV Tat action is unclear. There is no evidence to support direct binding of the Tat protein to the viral LTR sequences (Gdovin *et al.*, 1992, Neuveut *et al.*, 1993). Both of these studies did, however, report transactivation by the Tat protein. The transactivation potential of the MVV Tat protein is much reduced when compared to the HIV protein. MVV Tat can elevate LTR activity by between 5 and 30 fold; with the lowest transactivation observed using monocyte/macrophage cell lines (Neuveut *et al.*, 1993). The transactivation appears to occur primarily at the transcriptional level but there is also some evidence for a post transcriptional effect (Gdovin *et al.*, 1992).

One feature of the MVV LTR which distinguishes it from the HIV LTR is its high level of activity in the absence of Tat (Twu *et al.*, 1989). In comparison with the HIV U3 promoter the MVV sequences show a much greater activity. This difference is probably not due to the MVV LTR being a better initiator of pol II complexes. Rather it is probable that the complexes generated are more processive and do not require stabilisation. The HIV LTR has been described as an 'attenuated' promoter (reviewed by Jones & Peterlin 1994). The poor transcription from the HIV LTR is demonstrated when exogenous Tat is introduced into cells carrying endogenous Tat defective virus (Feinberg *et al.*, 1991). Only very low levels of virus could be detected from these cells in the absence of Tat. When exogenous Tat is introduced production of virus increases >30,000 fold. This effect is mediated via an increased efficiency of elongation of nascent transcripts, rather than effects on pol II initiation which is determined by U3 promoter sequences. The RNA pol II complexes formed on the HIV LTR are poorly processive and generate predominantly small RNA species that terminate shortly after initiation (Laspia *et al.*, 1989). The poor processivity of pol II complexes formed on the HIV LTR appears to be the cause of this accumulation of short transcripts rather than termination sites within the viral genome (Kessler & Mattews 1992). The Tat protein of HIV thus appears to be acting on pol II complexes formed by the LTR promoter sequences to elevate their efficiency or processivity of elongation. The sequence responsible for this targeting is the TAR sequence in the nascent RNA. As stated before deletion of *tat* or TAR abolishes viral replication. However, this action of Tat is not TAR dependent as deletion of the TAR in conjunction with targeting of the Tat protein directly to the HIV LTR using chimeric DNA binding constructs allows full Tat induction (Berkhout *et al.*, 1990, Kamine *et al.*, 1991, Southgate & Green 1991).

The HIV LTR can then be said to possess to distinct activities; a basal activity which initiates poorly processive pol II complexes at a high rate and a second Tat dependent activity which is required to generate fully processive complexes. The second, Tat dependent activity, relieves the block on elongation and processivity and allows the generation of high levels of full length transcripts. Both these activities, basal and Tat induced, are required for the viral lifecycle to proceed. Deletion of elements required for the basal activity of the LTR also prevents or attenuates viral replication (Leonard *et al.*, 1989, Ross *et al.*, 1991).

In MVV it appears that the LTR is capable of forming fully processive pol II complexes. Thus, together with the apparent absence of a TAR region it would appear that the MVV Tat

protein transactivates by a distinct mechanism to that seen with the HIV Tat protein. If MVV Tat is not binding to the promoter directly, as observed with a number of viral transactivators, or to an RNA TAR structure, how is it functioning? There are several examples of viral transactivators which function via interactions with cellular proteins rather than directly with sequences in the viral genome. Two of the better characterised examples are the HTLV Tax protein and the X protein (pX) of Hepatitis B virus (HBV). Like the MVV Tat protein neither of these transactivitors binds directly to the viral promoter or to any identified sequence in cellular genes. Thus they may represent better model systems than HIV Tat for asking how MVV Tat may be functioning.

The pX protein of HBV is capable of transactivating both viral and cellular genes (Twu & Robinson 1989, Hu *et al.*, 1990). Transactivation of MHC class I (Hu *et al.*, 1990) and class II (Zhou *et al.*, 1990) genes is observed together with the ß interferon promoter (Twu & Schloemer 1987). Viral promoters transactivated include HIV-1, SV-40 enhancer but not the MVV or HTLV-I LTRs (Twu & Robinson 1989). The pX protein has been shown to bind directly the cellular bZIP transcription factors CREB and ATF-2. This interaction alters the sequence recognition specificity of these factors, increasing their affinity for sequences within the HBV enhancer (Maguire *et al.*, 1991). The pX protein can also function through a number of other cellular factors including NF-kB, by which it up regulates the HIV LTR (Siddiqui *et al.*, 1989, Twu *et al.*, 1989), AP-1 and AP-2 (Seto *et al.*, 1990). The pX protein cannot transactivate through AP-3 or OCT recognition sites (Seto *et al.*, 1990) demonstrating that while a broad range of host factors are targets the effect is not general.

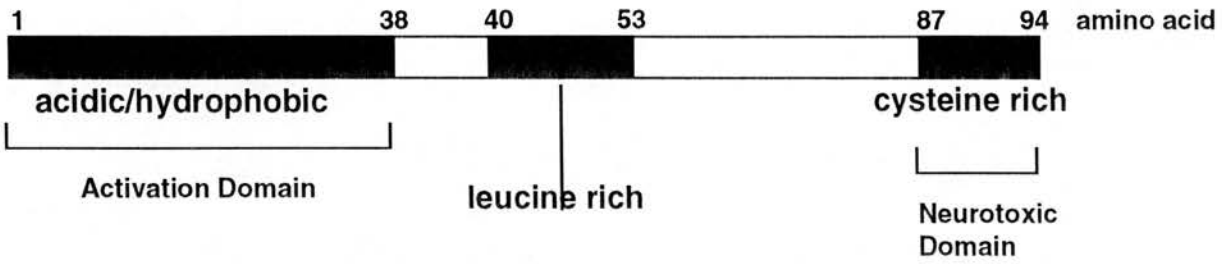
The HTLV Tax protein, although structurally distinct, targets similar cellular pathways. The Tax protein will bind to CREB and the closely related CREM protein, via a conserved amino acid motif adjacent to the DNA binding domain of these two proteins (Adya *et al.*, 1994, Suzuki *et al.*, 1993). This interaction is apparently restricted to these two bZIP proteins as no binding was observed to ATF-1 (Adya *et al.*, 1994). Again this protein/protein interaction modifies the DNA binding specificty of the transcription factors. The binding of the resultant protein complex to the promoter also serves to target the transcription activation domain of the Tax protein to the assembling pol II complex (Fujisawa *et al.*, 1991). Where concentrations of transcription factors are not limiting, Tax has been reported to enhance the dimerization of many bZIP transcription factors in addition to CREB (Wagner & Green 1993). A second distinct action of the Tax protein is to activate NF-κB by triggering its translocation to the nucleus (Leung & Nabel 1988). This function is mediated by a domain distinct to that involved in the

modulation of bZIP transcription factors (Smith & Greene 1990, Semmes & Jeang 1992). The mechanism of Tax induction of NF-κB is still unclear but it appears to target IκB, which holds NF-κB in the cytoplasm, rather than induction of the promoters of NF-κB subunits (Kanno *et al.*, 1994, Munoz *et al.*, 1994). Tax can interact with the Rel homolgy domain of the NF-κB proteins (Suzuki *et al.* 1994) although it is not known if this interaction alters the DNA binding specificity of the NF-κB complexes.

These two proteins, pX and Tax, belong to viruses from separate families and show no sequence homology. The pX protein does show homology with known serine protease inhibitors (Takada & Koike 1990). This may be of functional significance as high expression of the HIV LTR in cell lines (Fransozo *et al.*, 1994) has been shown to be dependent on the absence of a monocyte restricted family of serine proteases which modulate NF-κB activity. This highlights a possible mechanism for pX mediated elevation of NF-κB activities; pX may block the proteolytic degradation of functional NF-κB molecules.

This outline of the properties of the viral transactivators pX and Tax shows how different viral proteins can function by modifying the same cellular pathways at distinct points. These two proteins also highlight the domain structure possessed by a number of other viral transactivators where distinct functions in the transactivation process are mediated independently by separate domains. How does this relate to the MVV Tat protein? The exact mode of MVV Tat action is still unknown. Studies on the LTR of the 1514 molecular clone have, however, identified *cis* acting sequences within the LTR which are Tat responsive. This region of the 1514 LTR contains both a degenerate AP-4 site and a consensus AP-1 site. Transactivation can be mediated through either region but is optimal when both are present (Gdovin *et al.*, 1992, Neuveut *et al.*, 1993). The Tat protein will also transactivate heterologous promoters which bind AP-1 (Gdovin *et al.*, 1992). Attempts to show direct, or indirect (via cellular transcription factors), binding of Tat to the MVV LTR have failed (Neuveut *et al.*, 1993). It is still likely that the MVV Tat protein is functioning, at least in part, by being targeted to promoters. Functional analysis of Tat/Gal4 fusion proteins, which allows Tat to be introduced to a basal promoter via Gal 4 binding sites, revealed a potent transcription activation domain which shows homology to other acidic activators (Carruth *et al.*, 1994). The structure of the Tat protein of MVV strain 1514 is shown in Fig.1.2. This domain structure is conserved between all MVV isolates. Like other viral transactivators it possesses a number of distinct domains each of which may play a role in its regulation and function. The leucine rich domain has been

FIGURE 1.2 Structure of the MVV Tat Protein



The domain structure of the Tat protein of MVV is outlined above. Three domains have been identified; the acidic/hydrophobic, leucine rich and cysteine rich. The function of the various domains is discussed in the main text.

proposed as a means for the MVV Tat protein to target its activation domain to the promoter. Although this sequence is not extensive enough to function as a full leucine zipper as found in the bZIP transcription factors it could function to allow the Tat protein to hook itself onto bZIP dimers such as AP-1 (which is composed of the Fos and Jun proteins). This type of loose targeting interaction may explain the difficulty of demonstrating Tat interactions with promoter/transcription factor complexes. Studies on the full length Tat protein (Carruth *et al.*, 1994) revealed the possibility of negative regulatory elements within the Tat protein itself. These sequences would control the activation domain in the absence of the correct protein/protein interactions. Interaction with leucine zipper containing transcription factors (such as AP-1) may be required to generate the conformational changes required to 'unmask' the activation domain. Although conjecture, this potential mode of action is analogous to that seen in the Tax and pX transactivators. The observation that Tat can transactivate cellular genes, including *c-jun* (Neuveut *et al.*, 1993), suggests a general mode of action not primarily targeted to the viral LTR. Any promoter containing binding sites for the factors with which Tat interacts, or induces, will be a potential target for transactivation. In MVV where Tat *trans*-activation of the LTR is poor this effect on cellular genes may in fact be the function of the Tat protein. By maintaining the cell in an activated state the Tat protein can indirectly boost, and maintain, viral transcription.

The cysteine rich domain of Tat may also play a role in the function of MVV Tat. Peptides containing this domain are highly neurotoxic (Philippon *et al.*, 1994, Hayman *et al.*, 1993) and appear to interfere with signal transduction processes in neurones. So in common with other viral transactivators MVV Tat appears to posses several discrete functional domains which may to some degree function independently even though present within the same protein.

MVV is not the only lentivirus which shows a mechanism of Tat transactivation apparently independent of a TAR region. FIV which appears more closely related to the ruminant than primate lentiviruses (reviewed by Miyazawa *et al.*, 1994a) also appears to possess no TAR region and be only partially dependent on Tat (orf A) for viral replication. In all three of these viruses (MVV, CAEV and FIV) the Tat protein appears to be only a marginal transactivator of viral gene expression. In FIV, Tat deleted virus will grow in lymphoblastoid lines (Miyazawa *et al.*, 1993, Tomonaga *et al.*, 1993). However, Tat deleted virus grows poorly in primary feline T cell lines. The CAEV Tat protein has been reported as essential for *in vitro*

replication of the virus (Saltarelli *et al.*, 1993). Other workers have not been able to reproduce this inhibition and have suggested that Tat is not required for *in vitro* virus growth (A. Harmache personal communication). This, together with the prior discussion of the MVV protein, would suggest that the FIV, CAEV and MVV Tat proteins may be more important for *in vivo* rather than *in vitro* replication. The Tat proteins of these viruses may be involved in the modulation of normal cellular processes so upregulating, or enhancing, viral gene expression.

1.6 Control of MVV Transcription

1.6.1 Structure and Functional Domains of the MVV LTR

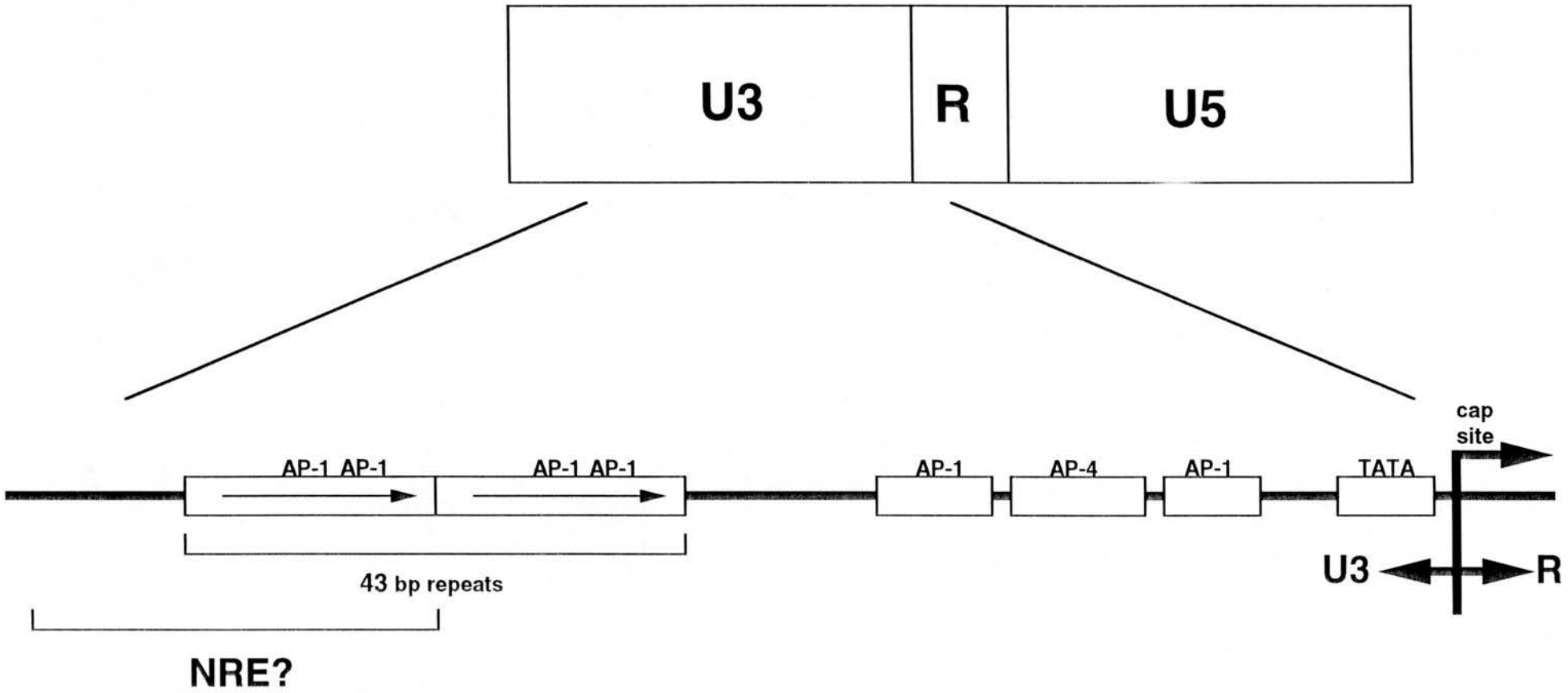
Lentiviral transcription is controlled by two interacting components; basal transcription, driven solely by the LTR sequences and Tat dependent transcription, which acts on the basal component to transactivate transcription. As discussed in Section 1.5.3 the basal component is unable to drive HIV-1 replication which is Tat dependent. In the non-primate lentiviruses MVV, CAEV and FIV there is experimental evidence to support the proposition that the basal component is competent to support viral replication in certain cell systems *in vitro* (A. Harmache personnal communication, Miyazawa *et al.*, 1993, Tomonaga *et al.*, 1993). This would suggest that these three viruses, unlike HIV, do not possess 'attenuated' promoters in the LTR which require the action of Tat to produce fully processive pol II complexes. In comparison with HIV there is still relatively little known about the regulatory sequences of the MVV LTR and the role of these various motifs in viral replication. In Fig.1.3 a schematic outline of the U3 promoter region of the MVV LTR is shown. All studies to date on the regulation of MVV transcription have been performed on the 1514 LTR. The other isolates have a similar but divergent structure and sequence (Chapter 3). A number of potential AP-1 sites have been proposed in the LTR. Of these only the TATA box proximal sequence matches the consensus sequence for an AP-1 site (Hess *et al.*, 1989, Faisst & Meyer 1992).

The LTR sequence shows strong promoter activity and position independent enhancer activity (Hess *et al.*, 1985). The 1514 LTR appears capable of driving a limited cell type specific expression of a reporter gene in transgenic animals (Small *et al.*, 1989, Clements *et al.*, 1994). In transgenic mice expressing the chloramphenicol acetyl transferase gene under the control of the LTR, expression is restricted to activated macrophages, lymphocytes,

FIGURE 1.3

Potential regulatory sequences within the 1514 LTR are shown. The positions of the various AP-1 like sequences and the AP-4 site are marked, as is the potential negative regulatory element (NRE).

Figure 1.3 Structure of the LTR of MVV molecular clone 1514



neurones and other cells in the CNS (Small *et al.*, 1989). Cellular activation appears to be a key trigger for the activation of transcription of the MVV LTR. Activation, *in vitro*, of monocytes or macrophages from these mice resulted in a further increase in transgene expression. Transgenic sheep expressing the MVV *env* gene under the control of the LTR only show expression of the Env protein in activated macrophages (Clements *et al.*, 1994). However, expression of the transgene *in vivo* has not yet been extensively studied in this system. They have shown that while *in vivo* there is no evidence for Env expression in fibroblasts, upon *in vitro* culture Env is expressed and syncitia form. This *in vitro* passage is likely to result in cellular activation, which would appear to be required for MVV LTR activity.

These data do not suggest that the LTR is responsible for the *in vivo* tissue tropism of the virus as expression, following activation, can be demonstarted in a variety of cell types. Rather, it appears that *in vivo* LTR activation may be favoured in activated macrophages due to the set of transcription factors produced as these cells differentiate and become activated. As in HIV it is likely that tissue tropism is determined by a stage prior to provirus formation (Schuitemaker *et al.*, 1993).

Transcription from the MVV LTR may be activated by serum or phorbol ester (Clements *et al.*, 1992). This activity has been mapped using deletion mutants of the 1514 LTR (Hess *et al.*, 1989, Clements *et al.*, 1992). Figure 1.4 shows an outline of the functional significance of various LTR sequences, as determined by deletion analysis (Hess *et al.*, 1989), and their effect on basal LTR activity. These studies identified several upstream regions which appeared to contribute to the LTR's activity. At the 3' end of the U3 region a negative regulatory element (NRE) was identified. Removal of this sequence resulted in an elevation of LTR activity. Deletion of sequences around the start of the first 43bp repeat resulted in a drop in activity suggesting the interaction of a positive regulatory factor with this site. Further deletions towards the TATA box resulted in a steady decline in activity with deletion of the AP-4 and TATA box proximal AP-1 site leaving an activity only just above background (Hess *et al.*, 1989).

As well as examining the contribution of sequences to LTR activity these workers also tested the constructs for their response to induction via phorbol ester. These results are summarised in Fig.1.5. The deletions were observed to have distinct effects on LTR induction. Removal of the NRE resulted in an enhanced response to stimulation. Deletion of the sequence near the 43bp repeat resulted in only a partial loss of the inducible response.

FIGURE 1.4

The effects on LTR activity, as measure by CAT assay, of 5' deletions are shown (from Hess *et al.*, 1989, Clements *et al.*, 1994). The deletions from the 5' end are indicated below an outline of the various transcription factor binding sites in the LTR.

FIGURE 1.4 Effects of Deletions on LTR Activity

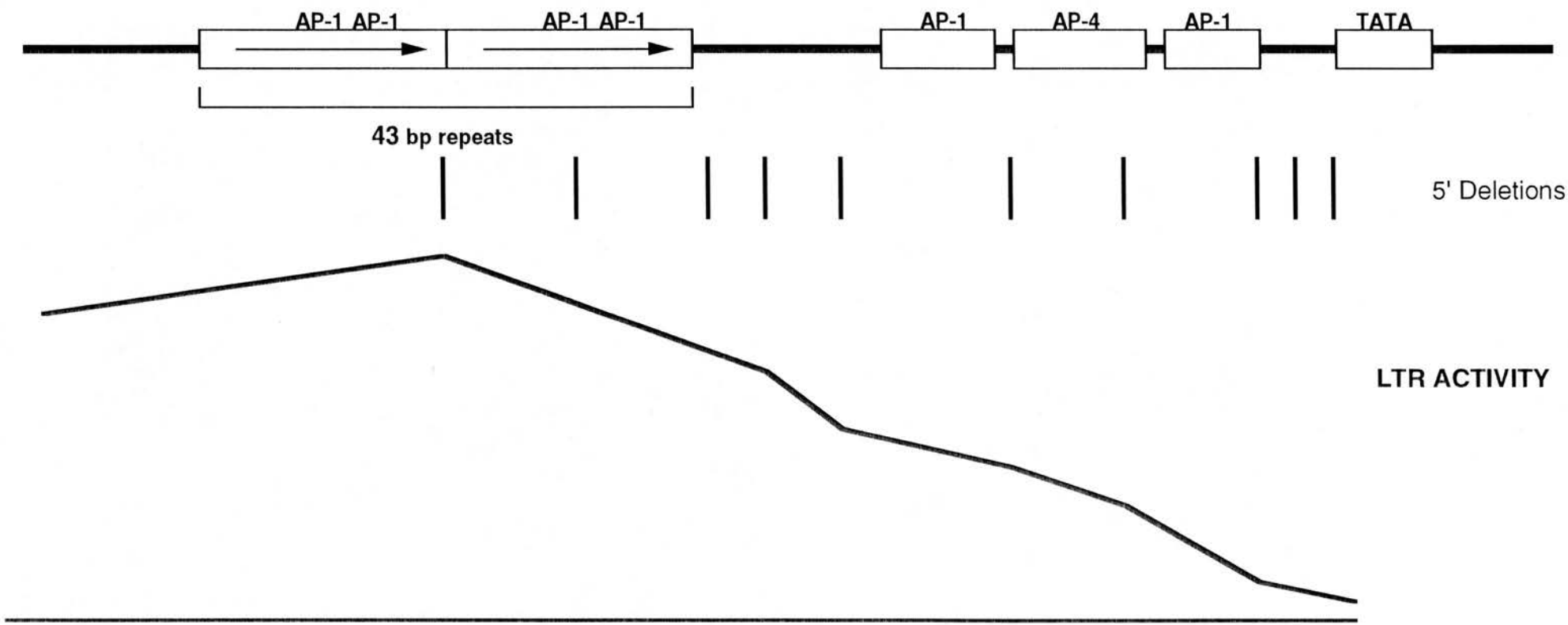
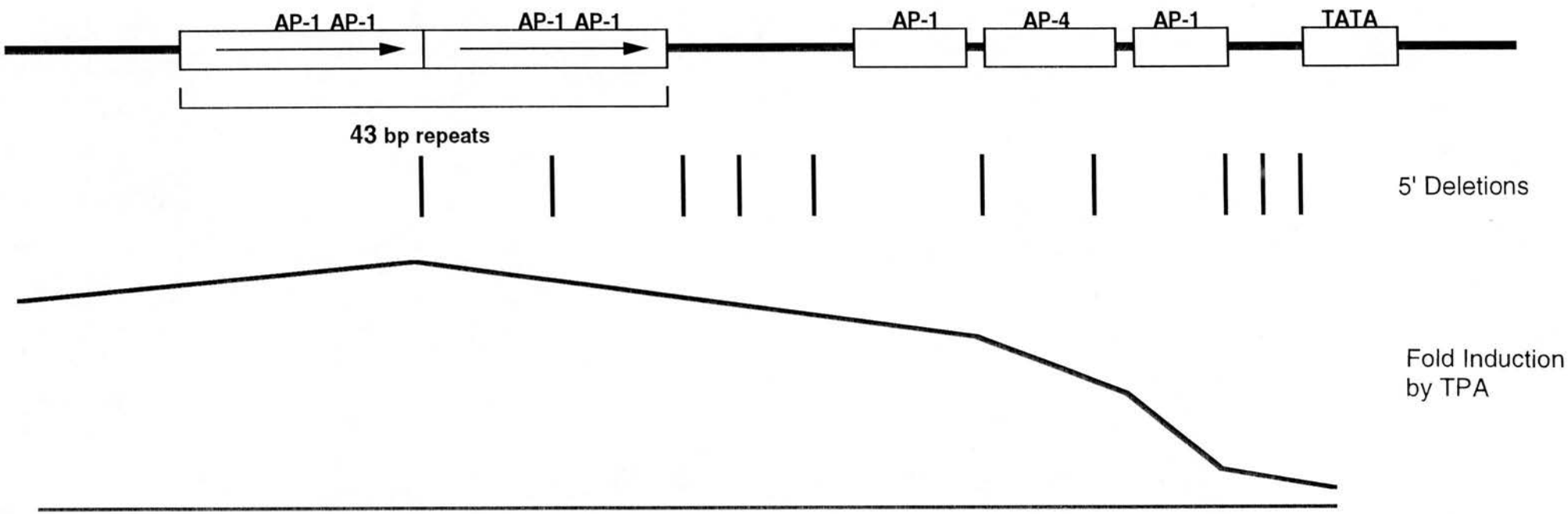


FIGURE 1.5

The inducible response of the 5' deletions is illustrated (Hess *et al.*, 1989, Clements *et al.*, 1994). The same deletions were examined by CAT assay for their response to TPA.

FIGURE 1.5 Effects of Deletions on TPA Induction



Induction appeared to be highly dependent on the AP-4 and TATA box proximal AP-1 site; removal of these sequences blocked induction. When the effect of these sequences in the context of the whole LTR was examined by linker scanner mutants this effect was confirmed (Hess *et al.*, 1989). These studies highlighted the interaction of various elements within the LTR and their apparently distinct functional roles.

Studies on the interaction of the cellular factors with the LTR revealed three protected regions (Gabuzda *et al.*, 1989). This data is shown in Fig.1.6. Of these footprints only one overlies the consensus sequence for a known transcription factor; the TATA box proximal consensus AP-1 site. Further studies on this sequence (Shih *et al.*, 1992) have demonstrated binding of a Fos and Jun containing complex (AP-1) to this sequence using extracts from the human U937 monocytic cell line. The footprint over this consensus site was observed to be phorbol ester inducible (Gabuzda *et al.*, 1989). No footprints were observed over the other AP-1 sites. This may reflect their non-consensus nature (Chapters 3 & 4).

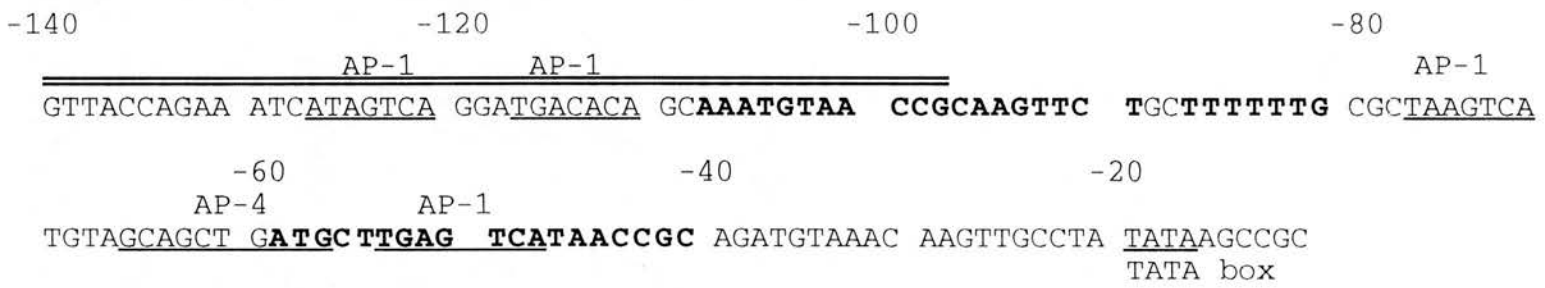
The nature of the factors binding to the other two sequences were not elucidated. The footprint over the coding strand from positions -87 to -81 appears homologous to a region footprinted in the HIV-1 LTR (Garcia *et al.*, 1987) and the HTLV-I LTR (Nyborg *et al.*, 1988) but the nature of this factor remains to be elucidated. Its presence in a number of lentiviral LTRs suggests that it may be playing an important role in the transcription processes of these viruses. The footprint over the -108 to -90 region appears to play an important role in the activity of the LTR. Deletion of this region results in a sharp drop in LTR activity (Hess *et al.*, 1989, Clements *et al.*, 1992, Fig.1.4.).

Studies on the regulation of MVV transcription have to date been limited to the 1514 virus. While all MVV strains share a common LTR structure, sequence variation between strains could potentially affect the binding of cellular factors and thus the activity of the LTR. This variation and its possible significance of this variation will be discussed in Section 1.6.3.

1.6.2 The AP-1 Complex

The transcription factor AP-1 is a dimer composed of members of the Fos and Jun families of proteins (Johnson & McKnight 1989). The genes which compose these protein families are called immediate early; cellular activation triggers their rapid upregulation. Although different proteins are regulated in distinct manners. Three members of the Jun family have

FIGURE 1.6. Footprinted Regions in the 1514 LTR



Sequence of the 1514 region footprinted in Gabuzda *et al.* (1989). The first 43bp repeat is double overlined. The potential AP-1 and AP-4 sites identified by these workers are underlined as is the TATA box sequence. The sequences observed to footprint using extracts from Phorbol ester induced U937 cells are in bold type. The footprint over positions -87 to -81 was observed only on the coding strand. The footprint over the TATA box proximal site extended from -41 to -57 on the coding strand and to -59 in the non-coding strand. The footprint from -108 to -90 covered the same region on both strands. Base numbering is from the RNA cap site.

been identified c-Jun, JunB and JunD (Ryder *et al.*, 1988, Ryseck *et al.*, 1988, Ryder *et al.*, 1989). These three proteins can form both homo and heterodimers (Schutte *et al.*, 1989). Dimerization is more efficient between the Fos and Jun families than within the Jun family. This more efficient dimerization leads to a higher affinity for AP-1 sites (Halazonetis *et al.*, 1988, Nakabeppu *et al.*, 1988, Jones 1990). The Fos family is composed of four characterised proteins; Fos, FosB, Fos-related-antigen-1 (Fra-1) and Fra-2 (Cohen & Curran 1988, Foletta *et al.*, 1994). Unlike the Jun family, these proteins do not form homodimers and must interact with a Jun family member in order to bind DNA. These proteins contain 3 main domains; a basic region which interacts with DNA, a leucine zipper involved in dimerisation (Sassone-Corsi *et al.*, 1988) and an acidic activation domain. At the transcriptional level these proteins are regulated according to cell lineage (Wilkinson *et al.*, 1989, Hilberg *et al.*, 1993, Foletta *et al.*, 1994), activation signals (Ryder & Nathans 1988, Saffen *et al.*, 1988, Redner *et al.*, 1992, Muegge *et al.*, 1993, Yoshida *et al.*, 1993), differentiation state (Matsui *et al.*, 1990, Hilberg *et al.*, 1993, Agamennon *et al.*, 1994) and cell cycle (Ryseck *et al.*, 1988,Kovary & Bravo 1991). In addition to this transcriptional regulation the protein products show distinct stabilities within cells (Carillo *et al.*, 1994, Gruda *et al.*, 1994, Kovary & Bravo 1991)

AP-1 is activity is controlled by several distinct post-translational mechanisms. This post translational regulation is demonstrated by the ability to induce AP-1 activity even in the presence of protein synthesis inhibitors (Farina *et al.*, 1993). Three distinct regulatory mechanisms have been identified; phosphorylation, redox state and interaction with regulatory proteins that block DNA binding activity. The involvement of a mechanism for cytoplasmic retention of AP-1 was identified following the observation that nuclear transport of Fos and Jun is not constitutive but requires continuing cell activation (Roux *et al.*, 1990). Two proteins have been identified which interact with AP-1. The IP-1 protein appears to interact with Fos and Jun via their leucine zipper domains (Auwerx & Sassone-Corsi 1991). IP-1 is inactivated by phosphorylation following Protein kinase A activation. Phosphorylation blocks the IP-1/AP-1 interaction so generating active AP-1 complexes. The Jif-1 protein appears to the homolog of an anti-oncogene, QM (Monteclaro & Vogt 1993). Unlike IP-1, Jif-1 does not interact with the leucine zipper and appears solely to bind and inhibit the Jun proteins.

A second post-translation regulatory system involves phosphorylation of the Jun and Fos proteins. The consequences of phosphorylation depend on which domain of the proteins is phosphorylated (Karin & Smeal 1992). Phosphorylation of Jun within the C-terminal

DNA binding domain leads to the repression of DNA binding activity (Boyle *et al.*, 1991). In contrast phosphorylation within the transactivation (N-terminal) domain has been proposed to stimulate the activation capacity of AP-1 (Binetry *et al.*, 1991, Pulverer *et al.*, 1991, Bannister *et al.*, 1994). Regulation of JunB and JunD by phosphorylation of sites in the C-termini has also been described (de Groot *et al.*, 1992, Nikolahahi *et al.*, 1993). The Fos protein is also regulated by phosphorylation (Ofir *et al.*, 1990). In this case phosphorylation is required for Fos repression of the *fos* gene. Although the hypophosphorylated form cannot repress *fos* transcription it can transactivate through a phorbol response element (AP-1 site). Phosphorylation of Fra-1 and Fra-2 has also been reported (Gruda *et al.*, 1994). For these proteins *in vitro* phosphorylation by MAP kinases was shown to stimulate their DNA binding activity. Both MAP kinases and Ha-Ras have been proposed as regulators of the Fos and Jun family proteins (Gruda *et al.*, 1994, Pulverer *et al.*, 1991, Binetry *et al.*, 1991). The phosphorylation of Jun also controls Fos stability (Papavassilou *et al.*, 1992). In the absence of Jun, or when bound to dephosphorylated Jun, the Fos protein was observed to be stabilised. However, interaction with phosphorylated Jun led to the rapid degradation of Fos. Phosphorylation of Jun and Fos appears to be dependent on the composition of the AP-1 dimer and on whether or not it is bound to DNA (Abate *et al.*, 1993). AP-1 is not unique in that phosphorylation affects its DNA binding activity. CREB, a bZIP transcription factor from the ATF-CREB family, is also modulated by phosphorylation (Nicols *et al.*, 1992). In this case aphosphorylated CREB can only bind symmetrical CRE sites (high affinity) but not asymmetric (low affinity) sites. In contrast, the phosphorylated form is capable of binding both sites efficiently. Phosphorylation is thus a common mechanism for the regulation of transcription factor activity.

The final mechanism identified for the regulation of AP-1 is the redox state of the Jun and Fos proteins (Abate *et al.*, 1990). Reduction of a cysteine residue in the DNA binding domains of Fos and Jun by a cellular protein results in an enhanced DNA binding activity (Abate *et al.*, 1990, Xanthoudakis *et al.*, 1992). The context of this cysteine residue also appears to be important in the redox regulation of Fos and Jun (Ng *et al.*, 1993). The relative contribution of these three regulatory pathways and their interaction in given cell types has yet to be determined.

Dimerisation occurs via the Leucine zipper domain, which is shared by two other transcription factor families; ATF-CREB and C/EBP. The shared domain allows dimerisation

between some of the proteins within these families (Ivashkiv *et al.*,1990, Hai & Curran 1991). The mechanism governing the dimerisation between families as well as within families is still poorly understood. Dimerisation between factors in the same, or different families appears to affect transcription in two ways; alteration of binding site selection and effects on transactivation. These effects can range from potentiation to inhibition of transactivation (Diamond *et al.*, 1990, Ivashkiv *et al.*, 1990, Ryseck & Bravo 1991, Suzuki *et al.*, 1991, Hsu *et al.*, 1994). The preferential recognition of half sites by the dimer components has not been studied in AP-1 but within the NF-kB family, which like AP-1 is composed of dimers, there is evidence that different members of the family have distinct half site preferences (Urban *et al.*, 1991). This means that the binding site will select which dimers in the population bind the site and also the complex will take on a specific orientation depending on the orientation of the site within the promoter. This orientation effect has been described in the HIV-1 LTR where NF-kB/Sp1 synergism is dependent on the orientation of the NF-kB sites (Perkins *et al.*, 1993).

The AP-1 complex preferentially recognises the sequence TGA $^{G}/_{C}T^{C}/_{A}A$ (Faisst & Meyer 1992). This consensus sequence for AP-1 binding represents an approximation to an ideal AP-1 site and its presence is not an absolute requirement for AP-1 binding (Mermod *et al.*, 1988, van Lint *et al.*, 1991, Jain *et al.*, 1992, Vallet & Osborne 1994). In addition to this core recognition site both flanking sequence and dimer composition will affect the affinity of any given AP-1 complex for any AP-1 site (Ryseck & Bravo 1991, Hadman *et al.*, 1993). Dimerization affects not only the affinity for DNA binding sites but also the consequences of binding. Dimer composition affects the ability of AP-1 to activate transcription (Ivashkiv *et al.*, 1990, Suzuki *et al.*, 1991) and the transforming/transcription activation activities of the Jun protein (Schutte *et al.*, 1988, Chiu *et al.*, 1989). In combination with c-Jun, Fra-2 causes suppression of transactivation caused by c-Jun alone, whereas JunD-Fra-2 complexes exhibit higher transactivation activity than JunD homodimers. In addition to these transactivation effects different dimers may cause distinct changes in DNA topology. Both Fos and Jun proteins are capable of bending DNA (Kerppola & Curran 1991a/b). However, the consequences for DNA topology vary on dimer composition; Jun-Fos and Jun-Jun dimers bend DNA in opposite directions. These could have two consequences; the bending induced by a given AP-1 dimer selects which further factors can interact with the DNA or vice-versa the promoter topology only allows the interaction of certain AP-1 complexes.

The role of an AP-1 site in the regulation of transcription can thus be seen to vary according to a number of conditions; cell lineage, the type of stimulus and context of the binding site within the promoter. Thus, the presence of a consensus site within a promoter sequence makes no predictions on the role of that site in the regulation of transcription from that promoter.

1.6.3 Other *cis*-regulatory elements

The sequences discussed in the previous section as being important for the regulation of MVV transcription are all positioned in the U3 region of the LTR upstream of the TATA box. It is possible that there are additional *cis*-acting regulatory sequences elsewhere in the genome.

In HIV-1 DNase hypersensitive sites (which represent regions containing an 'open' chromatin conformation, Section 1.7) are associated with both the LTR and the *pol* gene, which contains a genomic enhancer (Verdin *et al.*, 1990,Verdin *et al.*, 1991, van Lint *et al.*, 1994). This enhancer can be subdivided into two distinct regions one of which contains 3 binding sites for the transcripton factor AP-1 (van Lint *et al.*, 1991) and a second element which binds four distinct nuclear proteins (van Lint *et al.*, 1994). Both these elements are induced by phorbol esters. The presence of these inducible enhancer elements in HIV-1 may in part explain the observations that deletion of the NF-kB sites within the LTR do not block viral replication (Ross *et al.*, 1991, Kim *et al.*, 1993). While this may in part reflect redundancy in the LTR itself it may be that deletion of the inducible elements of the LTR may be compensated for by the inducible genomic enhancer. A further putative role for the enhancer is in the suppression of transcription from the 3' LTR. If this element preferentially acts on the 5' LTR then it may help downregulate the 3' sequence (Klaver & Berkhout 1994).

In addition to these sequences within the *pol* gene HIV-1 also contains a number of transcription factor binding sites downstream of the TATA box within the R/U5 region of the LTR (Kharroubi & Verdin 1994). Some of these sites, including Sp1 sites, appear to be constitutively occupied as observed by mapping of DNase1 hypersensitive sites (Verdin 1991).The LTR also contains a number of positioned nucleosomes (Verdin *et al.*,1993) one of which is positioned between the TATA box and these downstream elements. This nucleosome prevents DNase cutting of this region of the LTR in unactivated cells (Verdin

1991). However following activation of the cells and induction of transcription this site becomes hypersensitive (Verdin 1991, Verdin *et al.*, 1993). The modification of this nucleosome, which is positioned over the transcription initiation site, has been proposed as a key step in the activation of transcription from the LTR (Verdin *et al.*, 1993). The nucleosomes positioned elsewhere in the LTR do not appear to undergo such a marked modification following the activation of transcription.

In MVV the various sequences involved in the regulation of transcription have not yet been fully characterised. Using HIV-1 as a model system then it appears possible that there will be other regulatory sequences downstream of the TATA box, possibly including a genomic enhancer.

1.6.4 Components of Basal Transcription

As discussed in Section 1.5.3 lentiviral transcription may be divided into basal and Tat activated components. The basal component of MVV transcription is determined by the interaction of proteins bound to the regulatory sequences within the U3 promoter and the basal transcription machinery assembled on the TATA box. The MVV promoter has the classical structure of a pol II promoter; containing a TATA box and putative upstream sites for the binding of factors which can activate transcription. The TATA box (or core promoter element) is the site where the basal transcription complex assembles, including pol II, and determines the site of transcription initiation (reviewed in Buratowski 1994, Conway & Conway 1993). The upstream elements, plus any enhancers elsewhere in the genome of the virus (Section 1.6.3), determine the rate at which RNA polymerase II initiates new rounds of transcription. The transcription machinery can be divided into two classes of factors; *general* and *regulatory*. The general factors are involved in the formation of the basal complex at the TATA box and are assembled into the transcription complex in all or the majority of promoters. The regulatory factors are those which control the rate of transcription from a promoter and represent enhancer or promoter binding factors. The interaction of these to sets of factors is crucial for the regulation of transcription. The activity of the enhancer/promoter element can show dependence on the TATA box sequence (Wefald *et al.*, 1990). That is the exact complex of general factors is in part determined by the TATA box sequence and may be promoter specific. This complex may then respond differently to distinct upstream enhancer/promoter

elements, suggesting that the function of upstream regulatory elements can be determined by the basal complex assembled on the TATA box. This interaction of the general and regulatory components has also been shown in HIV where mutations to Sp1 sequences are compensated for by second site revertants in the TATA box which enhance the binding of TFIID and TFIIA (Kashanchi *et al.*, 1994). This results in a rescuing of these replication defective viruses via an augmentation of basal RNA synthesis. These TATA box mutations did not generate a more processive elongation complex and remained Tat inducible. These two examples illustrate the co-dependence of the basal/general transcription machinery and the regulatory factors.

The formation of the basal complex involves the assembly of over 20 proteins (Buratowski 1994, Conaway & Conaway 1993). The first step in this process is the binding TFIID to the TATA box. TFIID is composed of a number of polypeptides including the TATA box binding factor itself (TBP) and several other distinct associated proteins. A number of other proteins, called TAFs (Tightly Associated Factors), are also known to associate with TFIID (reviewed in Pugh & Tijan 1991, Gill & Tijan 1992). One important feature of TAFs is that while they are not required for basal transcription they are required for mediating regulated transcription.The factor TFIIA can interact with TFIID when bound to DNA (or in its absence) and appears to stabilise the interaction of TFIID with DNA. Once bound to DNA TFIID then recruits TFIIB (Hisatake *et al.*, 1993) which is capable of binding to pol II and other members of the basal complex, including TFIIF. A number of other factors are then recruited into this complex including TFIIH. Before transcription can procede fully the complex must be activated by the hydrolysis of ATP. This hydrolysis of ATP appears to be required, not for the initiation of transcription, but rather for the conversion of an initiation complex into an elongation complex (Goodrich & Tijan 1994). This hydrolysis step is apparently mediated by TFIIH and TFIIE, with TFIIH encoding the ATPase activity.

The rate of transcription is determined by the interaction of upstream factors with the basal complex described above. This interaction is mediated, in part, by a set of factors defined as co-activators (Berger *et al.*, 1990, Pugh & Tjian 1990, reviewed Tjian & Maniatis 1994). It appears that these co-activators are in some cases the previously identified TAF factors.

Although most activators of transcription characterised to date appear to interact with positive regulatory components of the basal complex there is another potential point of action.

One cellular protein, Dr1, has been identified as a negative regulator of the basal transcription machinery (Inostroza *et al.*, 1992). It appears that this factor blocks basal transcription via a direct interaction with TBP in the TFIID complex. This inhibitory step appears to be one target of the adenovirus transactivator E1A (Kraus *et al.*, 1994). E1A appears capable of blocking the association of Dr1 with TBP. This block is mediated by a Dr1 binding site in the N-terminus of the E1A protein. Thus, the presence of E1A in a promoter complex could block a Dr1 mediated inactivation of transcription. Dr1 is not the only example of a cellular negative regulator of basal transcription (Hahn 1993).

Activators of transcription have generally been described as targeting various components of the basal complex, including both the TFIID and TFIIB components (Tijan & Maniatis 1994). The regulatory factors which bind to promoter and enhancer elements contain activation domains which fall into several catagories; acidic (Hope & Struhl 1986), glutamine rich (Courey *et al.*, 1989) and proline rich (Mermod *et al.*, 1989). These activation domains mediate the interaction of transcription factors with the basal transcription machinery following their interaction with a promoter or enhancer. The interactions between transcription factors and the basal transcription complex may be either direct or indirect. The cofactors, which mediate indirect activation, were identified when it was shown that while purified TFIID could drive basal transcription it was unresponsive to activation by upstream Sp1 sites. Activation was seen when semi-purified TFIID was used in the same assays (Pugh & Tijan 1990, Smale *et al.*, 1990). Several TAF factors, or co-activators, have been described as interacting with transcriptional activators. The protein TAFII110 has been shown to bind the glutamine rich activation domains of Sp1(Hoey *et al.*, 1993) and a cell line with a temperature sensitive defect in TAFII250 was shown to be defective in Sp1 and VP16 activated transcription, but not basal transcription (Wang & Tijan 1994). Mutation in a region of the Sp1 activation domain, which contains alternating glutamine and hydrophobic residues, results in a decreased affinity for TAFII110 and a reduced ability of the protein to activate transcription (Gill *et al.*, 1994). While Sp1 has been shown to act primarily through theTFIID complex the viral transactivator VP16, which is targeted to promoters via its interaction with the cellular transcription factor Oct-1 (Kristie *et al.*, 1990), appears to interact with TFIID, TFIIB and TFIIH (Berger *et al.*, 1990, Goodrich *et al.*, 1993, Greenblatt *et al.*, 1994). Titration of the factor responsible for the activated transcription mediated by VP16 demonstrated that this factor is only involved in activated transcription and has no effect on the basal component (Berger *et al.*, 1990). The

protein TAFII40 binds both the acidic activation domain of VP16 and also the basal transcription factor TFIIB, suggesting a mechanism for enhanced recruitment of the transcription complex (Goodrich *et al.*, 1993). VP16 appears able to bind directly to TFIIB and enhance its a stable assembly into a pre-initiation complex (Lin & Green 1991). This binding of VP16 to TFIIB has been proposed to induce a conformational change in TFIIB. This conformational change enhances the binding of of TFIIB to other components of the basal complex which are recruited via their interaction with TFIIB, including TFIIF and pol II (Roberts & Green 1994).This interaction with TFIIB appears to be essential for the transcription activation mediated by VP-16 (Roberts *et al.*, 1993).

These two factors VP16, which contains an acidic activation domain (Hope & Struhl 1986), and Sp1, which contains a nonacidic glutamine rich activation domain (Courey *et al.*, 1989), appear to be activating transcription via two distinct, but overlapping, sets of co-activators. It should be noted that activation domains of the same type (i.e. acidic) may act through distinct sets of associated factors (Hahn 1993). This provides a model for how multiple distinct elements within a promoter can act synergistically to activate transcription, though it is still unclear how multiple sites for the same factor mediate this effect (Carey 1991,Tjian & Maniatis 1994). It has been proposed that activation of transcription is not solely dependent on the number of activation domains assembled on the promoter, but also the number of occupied factor binding sites (Oliviero & Struhl 1991). This would suggest a direct role for the DNA binding of transcription factors in the transcriptional activation process, perhaps by altering the conformatioon of the DNA in the promoter region.

It appears that transcription activators function by stimulating the rate of initiation by increasing the rate of assembly of the basal transcription complexes (Lin & Green 1991). It has been proposed that *in vivo* the interaction of the TBP component of TFIIB with the TATA box is a rate limiting step in transcription (Klein & Struhl 1994). These workers making use of TATA box mutants with upstream activator binding sites, proposed that activators function in part through alleviation of this block by enhancing the association of TFIID with the TATA box.

The MVV promoter (Section 1.6.1) contains a number of potential binding sites for activators of transcription. The relative importance of all these sequences has yet to be fully elucidated. In MVV this basal component appears to be highly efficient at driving transcription when compared to the HIV LTR and other cellular and viral promoters (Twu *et al.* ,1989). The data from transgenic animal experiments (Hess *et al.*, 1989, Clements 1994) suggests that *in*

vivo the block in transcription is mediated, at least in part, by cellular activation state. These observations support *in vitro* studies on the induction of AP-1 binding to the consensus site correlated with activation of the 1514 LTR (Gabuzda *et al.*, 1989, Shih *et al.*, 1992). Thus, MVV transcription appears to be tightly controlled by cellular activation state. The transcription factor AP-1 has been proposed as the major inducible regulator of MVV transcription (Clements *et al.*, 1992).

1.6.5 Tat Activated Transcription

As was mentioned in Section 1.5.3 the principle Tat responsive elements within the MVV viral genome appear to be sites where cellular transcription factors can bind the LTR (Gdovin & Clements 1992, Neuveut *et al.*, 1993). These studies identified the TATA box proximal AP-1 site as the main target for transactivation, but also suggested a role for the adjacent AP-4 site. It is still unclear how the MVV Tat protein mediates its transactivating effect, as it has not been possible to demonstrate association of Tat with cellular transcription factors (Neuveut *et al.*, 1993, J. Clements pers. comm.).

Studies on the transactivating potential of the MVV Tat protein consistently show only low levels of induction. This, together with the data of some workers showing Tat to be dispensable for *in vitro* replication, has led to the suggestion that the function of the MVV Tat protein is tightly regulated and that its main target may not be the viral LTR itself but rather the cellular genes for transcription factors which are involved in the control of transcription from the LTR, such as *c-fos* and *c-jun* (Neuveut *et al.*, 1993, Carruth *et al.*, 1994). The exact role of Tat in the transactivation of cellular genes remains unclear since some workers have shown induction of both *c-fos* and *c-jun* (J. Clemence personnal communication), while others have only observed induction of *c-jun* with no effect on *c-fos* (Neuveut *et al.*, 1993).

The work of Carruth *et al.* (1994) which mapped activation domains on Tat using Gal4/Tat fusion proteins which could be directly targeted to a test promoter revealed the acidic/hydrophobic region to be a potent activator of transcription (Fig.1.2.). In addition to the identification of this activation domain two regions were identified which possessed a potential negative regulatory function. One contained in amino acids 1-13, the other upstream of amino acid 38, potentially in the leucine rich domain. These workers proposed a mechanism for Tat action in which the free Tat protein has a minimal transactivating activity due to the action of the

negative regulatory elements. Following interaction with a cellular factor, perhaps Fos or Jun via the leucine rich domain, the activation domain is unmasked and the Tat protein transactivates the promoter it is targeted to via the cellular factor. These workers also used VP16 and HIV-1 Tat in competition experiments to 'squelch' MVV Tat transactivation. Overexpression of either of these two proteins interfered with MVV Tat transactivation suggesting that these three transactivators interact with an overlapping set of co-activator, or TAF, proteins (Section 1.6.2). The action of HIV Tat suggests that a post initiation activation may function in some cases to enhance elongation. HIV Tat has been shown to interact with the TFIID component of the basal complex (Kaschanchi *et al.*, 1994). This interaction would explain TAR independent activation of the HIV LTR by Tat (Harrich *et al.*,1990), suggesting the role of the TAR region may be primarily to target the Tat protein to the promoter and basal transcription complex. Thus multiple elements in the promoter can co-operate to enhance the rate of transcription.

The role of the MVV Tat in viral replication, *in vitro* and *in vivo*, remains unclear. The cellular factors with which it interacts remain to be fully characterised as does the set of cellular genes susceptible to transactivation and the significance of this process to viral replication and disease pathology.

1.7 Promoter Architecture

A number of distinct geographical isolates of MVV have been identified and sequenced (Section 1.2). The LTR sequences of these viruses shows variation both between and within isolates (Chapter 3). Sequence variation can potentially result in the alteration of transcription factor binding to the promoter due to the deletion of existing sites or creation of new sites via base changes in the sequence.

As well as this direct effect, alterations in transcription factor sites, or their position relative to other sites in the promoter, may have a more subtle effect on promoter activity. This effect is believed to be due to the 'stereospecific' complex which assembles on the promoter (reviewed in Tjian & Maniatis 1994). It would appear that in a number of promoter and enhancer systems transcriptional control is not solely dependent on the transcription factors interacting with the basal machinery but also the 3-dimensional organisation of these factors. The formation of these stereospecific complexes depends on the looping and bending of DNA to

allow the complex to form in the correct conformation. A number of transcription factors have been demonstrated to mediate the bending of DNA. This bending of DNA may result in either positive (Giese *et al.*, 1992, Giese & Grosschell 1993) or negative (Natesan & Gilman 1993) regulation of gene transcription.

DNA looping (reviewed in Schief 1992) has been proposed as the mechanism by which distant regulatory elements (e.g. enhancers) can mediate their effect. In the case of the TCRα gene the enhancer is located 4.5kb downstream of the most distal Cα exon (Ho *et al.*, 1989, Winoto & Baltimore 1989). This highlights the large distances over which these elements can operate. Changes in DNA structure would allow direct contact of these distal elements with basal transcription machinery assembled on the promoter. Protein induced changes in DNA structure may thus result in stereospecific complex formation at enhancers and promoters and also mediate the looping of DNA required to bring these elements together. There is still little direct evidence to date to prove that DNA looping does indeed bring promoter and enhancer elements into close proximity. Indirect evidence comes from several sources including prokaryotic systems (Schief 1992) and the ability of certain transcription factors, such as Sp1, to mediate DNA looping *in vitro* (Su *et al.*, 1991). These systems provide a mechanism by which distal elements can be brought together and the interaction stabilised. Further circumstantial evidence for DNA looping comes from the 'transvection' phenomenon in Drosophila.

Transvection occurs between certain combinations of mutant alleles and shows dependence on chromosome pairing or allele proximity (Judd 1988, Geyer *et al.*, 1990). The phenomenon involves the *trans* action of what are normally *cis* acting regulatory sequences. The *trans* action of the enhancer sequences, which allows them to drive the transcription of the other allele is dependent on the on the *cis* promoter sequences of the enhancer being inactive (Geyer *et al.*, 1990). This would suggest that under normal conditions the *cis* interaction is strongly favoured. Identification of the factors mediating the *trans* effect have shown that sequence specific transcription factors are responsible for mediating the effect (Biggin *et al.*, 1988). The phenomenon of transvection would appear to require DNA looping to allow regulatory elements on different chromosomes to interact.

The effect of transcription factors on local DNA structure is better understood. DNA bending *in vitro* has been observed using a number of different transcription factors; including YY1, LEF-1 and SRY (Giese *et al.*, 1992, Natesan & Gilman 1993). These three sequence

specific factors all recognise DNA via a high mobility group (HMG) domain. The bending of DNA induced by the factors YY1 and LEF1 has been described as essential for the function of the whole element in which they are present (Giese *et al.*, 1992, Carlsson *et al.*, 1993, Natesan & Gilman 1993). In the TCRα enhancer the minimal element contains a binding site for the LEF-1 transcription factor (Travis *et al.*, 1991, Waterman *et al.*, 1991) which is flanked by a site for CREB family members (Winoto & Baltimore 1989, Ho & Leiden 1990a) and an Ets site (Ho *et al.*, 1990). The LEF-1 factor is only active in the context of both these other two sites. Deletion of either non-LEF-1 site blocks enhancer function and sites for the LEF-1 factor alone do not activate transcription (Waterman & Jones 1990, Travis *et al.*, 1991, Carlsson *et al.*, 1993). This has led to the suggestion that HMG domain promoters are 'architectural' factors (Giese *et al.*, 1992, reviewed in Grosschedl *et al.*, 1994). While this is probably true in part, it is probably an over simplification. In addition to inducing DNA bending the LEF-1 factor also possesses a potent activation domain which may also be required for its function (Giese & Grosschedl 1993, Carlsson *et al.*, 1993).

The factors YY1 and SRY also appear to be regulating transcription via alterations in DNA architecture (Natesan & Gilman 1993, Cohen *et al.*, 1994). In the *c-fos* promoter the action of the YY1 factor is to repress transcription. This does not appear to be an intrinsic activity of the YY1 factor but rather is an effect of its binding site positioning within the promoter (Natesan & Gilman 1993). Alteration of the YY1 recognition by reversing its orientation within the promoter results in enhanced transcription from the promoter. This would suggest that in the normal context the factor is functioning to hold the promoter in a conformation which represses transcription. A similar mode of action has been proposed for the SRY factor in control of the *c-fra-1* promoter (Cohen *et al.*,1994), with the difference that in the normal promoter SRY activates *c-fra-1* transcription.

A further example of the formation of stereospecific complex formation at a promoter is provided by the IFN-ß promoter. The viral specific induction of transcription depends on a precisely assembled complex of transcription factors (Thanos & Maniatis 1992, Du *et al.*, 1993, Tjian & Maniatis 1994). This promoter contains regulatory elements which bind NF-kB, IRF-1 and ATF-2-c-Jun dimers (Du & Maniatis 1992). In addition to these factors there are 3 sites which specifically interact with the HMGI(Y) protein. The HMGI(Y) protein does not contain an HMG domain, but does contain a highly charged basic sequence which is required for its interaction with AT rich duplex DNA. These HMGI(Y) sites are believed to play a structural role

due to the ability of the protein to bend DNA. The interaction of HMGI(Y) with the promoter is required for the binding of NF-kB to the promoter (Thanos & Maniatis 1992). HMGI(Y) interacts with a site between the two halves of the NF-kB site. The induced bend allows the binding of NF-kB. The two factors appear able to bind simultaneously to the site, apparently because the HMGI(Y) protein is interacting through the minor groove. The interaction of HMGI(Y) with the promoter is not solely to make available DNA binding sites for other factors. The protein also appears to play a direct role in the protein-protein contacts taking place at the promoter even though it does not possess an activation domain itself (Du *et al.*, 1993). The bending of DNA by transcription factors is a general phenomenon not restricted to a single family. It is also seen with the 'forkhead' proteins and the AP-1 complex (Kerrpola & Curran 1991a/b, Pierrou *et al.*, 1994).

All the examples described so far relate primarily to protein induced alterations in DNA conformation. The structure of DNA can also effect the interaction of proteins bound to a promoter. In this case the spacing of elements can affect their interaction. This effect has been demonstrated in a number of systems (Cohen & Meselson 1986, Khalili *et al.*, 1986, Takahashi *et al.*, 1986, Comb *et al.*, 1988, Ondek *et al.*,1988, Schatz & Chatton 1990). Each turn of the DNA helix requires 10.5bp. Inserts between elements that represent half turns of the helix result in a reduction of transcriptional activity, presumably by blocking the protein-protein interactions of DNA binding factors. In the *Drosophila* Hsp70 promoter a periodic effect is seen for inserts up to 800bp (Cohen & Meselson 1986). Using the SV40 early promoter enhancer a similar effect was observed, though, the periodicity was lost with inserts over 125bp (Schatz & Chatton 1990). The limiting of the periodic effect to inserts below a certain length would suggest that when the spacing reaches a certain level the protein-protein interactions involved in activating transcription can bend the DNA to allow stable interactions. It is important to note that spacing may have effects beyond simply rotating previously adjacent sites so blocking their interaction (Huang & Jeang 1993, Khalili *et al.*, 1986). A more subtle effect may also be at work. In the HIV-1 LTR an increase in the TATA box Sp-1 sites spacing of between 18 to 101 nucleotides results in either a partially or wholly replication defective virus (Huang & Jeang 1993). For short inserts between the TATA box and Sp1 sites (10-40bp) an increase in basal transcription was observed. However, these inserts blocked Tat transactivation so preventing viral replication. For longer inserts both basal and Tat activated transcription were reduced. The effect of these insertions does not appear to be rotational but

rather would seem to be preventing the normal interaction of Sp1 with the basal transcriptional machinery in some other way.

The spacing effect has also been described in enhancer elements. In the SV-40 system the enhancer has two levels of organisation both of which show redundancy (Ondek *et al.* 1988). Enhancers are made up of multiple binding sites for cellular factors. Groups of these sites ('enhancons'), can be duplicated or interchanged to create new enhancer elements. Even though the full enhancer element is position independent, the spacing of enhancons within the enhancer is crucial for proper function. This spacing effect is also observed in the HIV LTR where NF-κB and Sp1 interact to drive transcription (Perkins *et al.*, 1993). The activity of this enhancer is dependent on both site orientation and spacing. Any increase in spacing between the sites reduces the enhancers activity. This effect does not appear to be due to a steric effect caused by rotation of sites around the DNA helix but rather these two factors, Sp1 and NF-κB, must interact directly to activate transcription. This interaction is mediated by the DNA binding domains, or a closely positioned sequence, rather than the activation domains of the proteins (Perkins *et al.*, 1994). Thus, any increase in site spacing will interfere with the protein/protein contacts which are required for enhancer function. This also explains the orientation effect as NF-kB site recognition is determined by the subunits of the dimer (Urban *et al.*, 1991). Since the Sp1/NF-κB interaction is dependent on the NF-κB unit involved the orientation of the NF-kB site will determine which component is positioned to interact with Sp-1. Co-operative interaction between NF-κB and Sp1 is also dependent on which member of the Sp family is bound to the GC boxes (Sp1 sites) in the HIV-1 LTR (Majello *et al.*, 1994). Only Sp1 co-operates with NF-κB to activate gene expression; neither Sp3 nor Sp4 is capable of co-operating with NF-κB. This interaction was mapped to the A *trans*-acting (activation) domain of Sp1. These two studies (Majello *et al.*, 1994, Perkins *et al.*, 1994) indicate the complex interactions, involving distinct protein domains, occuring between transcription factors at a promoter.

The formation of a stereospecific complex would thus appear to depend on two interacting components; the transcription factors bound to the DNA and the DNA structure of the promoter itself which will affect what interactions are possible. Before going on to discuss the evidence for promoter variation being significant in the control of gene transcription it is important to mention one further factor which may effect gene transcription; chromatin structure.

Within the nucleus DNA is complexed into a nucleoprotein structure known as chromatin (reviewed in Paranjape *et al.* 1994). Chromatin consisits of roughly a 2:1 mass ratio of protein to DNA and a 1:1 mass ratio of histones to DNA, the other protein components include the HMG proteins. Within the nucleus the chromatin structure is dynamic and is capable of moving between a number of states. As a generalisation chromatin can be divided into two catagories: euchromatin and heterochromatin. Within euchromatic regions the DNA is believed to be available for transcription; either constitutively or following the correct cellular activation events. Heterochromatin is more tightly condensed than euchromatin and insertion of euchromatic genes into heterochromatic regions results in their silencing. However, in *Drosophila* a number of genes are known to be present within heterochromatic regions (Gatti 1992). These genes appear to be specificly transcribed late in the cell cycle suggesting that their presence within heterochromatic regions may play a role in their regulation.

Following integration the retroviral genome is likely to behave in a manner similar to any cellular euchromatic gene. As discussed in section 1.4.1 the integration machinery appears to target the viral genome to regions of the cellular genome where transcription is active. Within euchromatic regions the chromatin is present in a 'beads-on-a-string' structure. This is the basic structure of chromatin, condensation of which leads to the compaction of the DNA and the formation of heterochromatin. Each 'bead' represents a nucleosome which contains DNA wound twice around the core histone octamer and an associated histone H1 which is not part of the core particle (Noll & Kornberg 1977). Following a short nuclease digestion a 'ladder' of DNA fragments is produced containing DNA from different numbers of nucleosomes. The typical length of DNA associated with each nucleosome is 180-210bp. Following a more extended nuclease digestion a 146bp ladder is generated. This represents the DNA which is associated with the core particle (Simpson 1978) and is the DNA bound to the histone octamer. The core particle does not contain histone H1. Histone H1, or linker histone, is thought to interact with the DNA between each core particle. Histone H1 is believed to play a role in the compaction of DNA and repression of transcription (Nacheva *et al.* 1989, Kamakaka & Thomas 1990, Bresnick *et al.* 1992) although its full role remains to be defined.

In addition to their structural role in the organisation of chromatin, nucleosomes may also have a more subtle role in the regulation of transcription. In a number of promoters it is clear that nucleosomes are 'positioned' *in vivo* (Benezra *et al.* 1986, Thomas & Elgin 1988,

Buckler *et al.* 1991, McPherson *et al.* 1993). The positioning effect is thought to be affected by the DNA sequence at the promoter. In general the minor groove of A/T rich sequences faces towards the histone octamer and the minor groove of G/C rich sequences away from the octamer (Satchwell *et al.* 1986, Satchwell & Travers 1989). Alternating A/T rich, G/C rich sequences result in a curved template which preferentially associates with the core histone octamer *in vitro*. This positioning effect has lead to the suggestion that nucleosomes may inhibit transcription by blocking the recognition sites for transcription factors (Archer *et al.* 1992). Alterations in chromatin structure at promoters can be seen by mapping DNase hypersensitive sites (Mueller & Wold 1989, Pfeifer & Riggs 1991, Verdin 1991). This indicates that remodeling of nucleosomes appears to accompany the activation of transcription. There is evidence to support the idea that while remodeling occurs the promoter is not truly nucleosome free. Thus nucleosomes at promoters are remodelled rather than fully displaced. The resulting alterations may generate a looser association of the histones with the DNA allowing fuller access for transcription factors. Chemical crosslinking studies using the *Drosophila* hsp70 genes indicate that core histones remain associated with the promoter following activation (Solomon *et al.* 1988, Nacheva *et al.* 1989). In the MMTV promoter it would appear that the induced DNase hypersensitive site is histone H1 depleted but remains associated with the core octamer (Bresnick *et al.* 1992). The maintenance of nucleosomes is also seen in the albumin enhancer (McPherson *et al.* 1993). This enhancer contains an array of precisely positioned transcription factors and nucleosomes. *In vitro* studies with artificially positioned nucleosomes have suggested that under low ionic strength conditions histone H1 reduces the mobility of nucleosome cores (Penning *et al.*, 1994). This suggests that the depletion of H1 at promoters is required to allow nucleosome remodeling, although whether this is cause or effect is unclear. The association of Histone H1 with the nucleosome core may have a differential effecton transcription factors. The basis of this interference remains to be determined (Juan *et al.*, 1994)

Nucleosome positioning can also affect the affinity of a transcription factor for its binding site (Li & Wrange 1993). Depending on the positioning of the nucleosome on the DNA sequence the GR/GRE (Glucocorticoid Receptor/ Element) affinity is seen to vary. This effect is not due to alterations in DNA sequence which affect positioning of the nucleosome on the DNA sequence, but simply requires the DNA to be bound to the core octamer. A similar effect is seen with Sp1 (Baiyong *et al.*, 1994). Like the GR, Sp1 is capable of binding its

recognition site even when it is bound to a histone octamer. The interaction of Sp1 with its binding site is 10 fold lower when the site is associated with a nucleosome core but binding does not result in a dissociation of the DNA-histone complex. Rather a ternary Sp1 nucleosome complex is formed. The interaction of Sp1 with a nucleosome associated site appears to be mediated solely via the DNA binding domain and appears not to require Sp1-histone interactions. These two examples, GR and Sp1, although carried out in *in vitro* systems suggest that nucleosomes need not be displaced to allow interactions between transcription factors and their recognition sites. Indeed *in vivo* studies on the albumin enhancer (McPherson *et al.*, 1993) suggest that such ternary complexes are also formed in the natural context.

The examples illustrated above show that a simple model of nucleosome involvement in transcription, where nucleosomes inhibit transcription by covering transcription factor sites, is an over simplification. These positioning effects suggests that there may be a combined interaction between transcription factors and positioned nucleosomes for enhancer and promoter function (Li & Wrange 1993, McPherson *et al.*, 1993, Lewin 1994). One effect which could be mediated by core nucleosomes which remain bound to a promoter is on closely positioned transcription factor sites. On a linear DNA molecule two closely positioned transcription factor sites may interfere due to the inability of both factors to bind. However if the same template is now bent, around a histone or via some other set of factors, both sites could potentially become available. While the full role of nucleosomes in the transcription process remains unclear it is possible that they may be mediating a number of distinct effects. One promoter which appears to show this complex interaction between transcription factors and nucleosomes is the HIV-1 LTR (Verdin *et al.*, 1993).

This section has highlighted the importance of 'stereospecifc' complexes at promoter elements. This appears to involve a number of interacting processess; histone positioning and alteration of DNA conformation, alteration of histone DNA interactions by sequence specific DNA binding factors and the role of non-histone DNA binding proteins such as HMGI(Y). Promoters cannot then be thought of as a series of discrete independent domains. Rather alteration of promoter structure due to changes in DNA sequence is likely to affect a number of complex interactions.

1.8 LTR Variation - Possible Effects on Viral Replication

The retroviral LTR contains the sequences which direct the transcription of viral genes in the correct cellular environment. As these sequences play a pivotal role in the viral lifecycle does variation in this region of the genome play a role in viral pathogenesis?

In the murine type C retroviruses LTR variation appears to play an important part in the control of viral pathogenicity (Golemis *et al.*, 1990). In a number of these viruses pathogenic phenotypes have been mapped to the enhancer/promoter element of the LTR. Simply exchanging a promoter/enhancer element can convert a non-oncogenic virus into an oncogenic form (Celander & Haseltine 1984). These two LTRs differ in their promoter activity and the cellular factors which bind to them (Boral *et al.*, 1989). Similar observations have been made in avian retroviruses where defects in endogenous viruses, which block expression, map to the LTR (Cullen *et al.*, 1983). Activiation of endogenous virus also appeared to be associated, in some cases, with the aquisition, via recombination, of new LTR sequences from exogenous virus (Tsichlis & Coffin 1979). This work on simple retroviruses would suggest that variation in promoter/enhancer sequences can lead to alterations in virus pathogenicity. Given the high rate of mutation in retroviruses (reviewed in Katz & Skalka 1990) there are likely to be a large number of different virus quasispecies with differing transcription rates available for selection *in vivo*.

The possible significance of promoter variation has also been studied in HIV-1. Promoter variation may manifest itself in one of two ways. Firstly, a relative enhancement or reduction of transcription when compared to other quasispecies within the infected individual and secondly alterations in promoter sequence which result in an altered tissue tropism of the virus; mutations which lead to the aquisition of sites which bind transcription factors specific to a certain cell type. In HIV-1 it is possible to isolate viral strains with distinct cell tropisms; ie T-cell, monocyte/macrophage, glial cell. Studies on the mechanism of this tropism have shown that the product of the *env* gene, gp120, is responsible (Sattentau *et al.*, 1988, Cheng-Meyer *et al.*, 1990, Shiola *et al.*, 1991, Westervelt *et al.*, 1992, Sharpless *et al.*, 1992). Using transfection of molecular clones for T-cell or macrophage tropic cell lines Schuitemaker *et al.* (1993) demonstrated that all these viruses could productively infect T-cells or monocyte-derived-macrophages (MDMs). However, when infection was performed by cell free virus the expected cell tropism was observed. These experiments would suggest that the LTR does

not itself control cell tropism. There is, however, evidence to suggest that the viral LTR may become cell type adapted after tropism has been determined (Corboy *et al.*, 1992). These workers generated transgenic mice carrying a reporter gene under the control of the HIV-1 LTR. The LTRs used came from molecular clones derived from either the CNS or T cells of an AIDS patient. Only the two LTRs from CNS derived virus were able to direct expression of the reporter gene in the CNS. The LTR derived from the T cell tropic virus showed no CNS expression. This data would suggest that LTR adaptation to the cell type which the virus is infecting may be occuring in HIV-1.

The second aspect of LTR promoter variation is the effect on transcription rate and growth rate and the significance, if any, of this to the infectious process *in vivo*. One recurring problem with this type of study is the source of material from which the LTR sequences are derived. Viral stocks maintained *in vitro* by passage in cell lines, or PBMs, undergo selection for this culture environment (Delassus *et al.*, 1991, Delassus *et al.*, 1992). These studies suggest that sequence alterations occur even following the replication involved in explant culture. Several studies have been performed to test whether the transcriptional activity of the LTR, and Tat transactivation, are selected to favour higher or lower transcription rates *in vivo*. Longitudinal studies using PCR amplified material derived from PBMs without *in vitro* culture appear to show that while variation in the both rate of basal transcription and degree of Tat transactivation does occur there appears to be no selection, as both high and low activity quasi-species occur simultaneously and fluctuate apparently at random (Delassus *et al.*, 1991, Michael *et al.*, 1994, Delassus *et al.*, 1992). Variations in the LTR promoter sequences include duplication of transcription factor sites and also point mutations which affect the binding of transcription factors (Michael *et al.*, 1994). In the studies described above only transcription rate and Tat transactivation was examined, not the rate of replication of the viruses carrying these variable LTR sequences.

Two groups have suggested that growth differences between HIV-1 molecular clones may be due to LTR sequence variation (Englund *et al.*, 1991, Koken *et al.*, 1992). In both these cases the molecular clones where derived from viral isolates using explant culture, so the degree of *in vitro* selection had been kept to a minimum. Koken *et al.* (1992) reported a 5-10% growth increase apparently due to the insertion of an additional Sp1 site. Englund *et al.* (1991) examined 3 molecular clones derived form the same individual. One of the viruses, which carried a deletion of an NF-kB site showed a lower transcription rate and a reduction in

viral replication and cytopathic effect. In SIV comparison of viral strains which induce acute and non-acute disease has identified a duplication within the LTR which inserts an extra NF-κB site (Dollard *et al.*, 1994). This duplication elevates LTR activity and increases the sensitivity of the LTR to cell activation. These workers generated virus which was isogenic for all regions other than the LTR and tested for any growth differences. From this experiment it was seen that the virus with the LTR derived from the 'acute' virus showed enhanced growth kinetics. It remains to be fully demonstrated that these LTR differences do indeed account for the differences in growth rates and pathology of these two viral strains.

The conclusion from the various studies outlined above is that there does not appear to be selection of viral promoters with a specific activity during infection. Even during the late stage of disease and progression to AIDS there does not appear to be selection of higher transcription rates (Delassus *et al.*, 1992, Nagushummugan *et al.*, 1992) although faster replicating viral forms are present at the end stages of disease. Despite this it appears that *in vivo* a large number of viral quasispecies with different transcription rates are tolerated. One aspect of viral replication which may be most dramatically affected by transcription rate is the conversion from early to late gene expression. In HIV this transition is under the control of the Tat and Rev proteins. Elevation of transcription via Tat leads to a positive feedback on the production of the small spliced RNAs. The accumulation of the products of these RNAs, most importantly Rev, alters the splicing pattern and switches on the expression of the structural genes. Following integration of the provirus the transition from a silent, latent state, to the production of the early proteins will be under the sole control of the LTR sequences. Once the threshold level of Tat is reached then it may transactivate viral and cellular genes. However, the absolute rate of transcription from the LTR will limit the accumulation of Tat and Rev; the weaker the promoter the longer it will take for these proteins to accumulate to the critical levels to trigger expression of the structural genes. It is possible that only this early stage of the viral lifecycle could be limited by the LTR transcription rate and that the later stages of infection are limited at other steps. While the transactivating proteins of MVV and HIV appear to be functioning by distinct pathways the viruses pass through similar steps in the infected cell.

1.9 AIMS

The general aim of this work was to attempt to gather more information on the regulation of MVV transcription. To date all the studies examining the control of transcription in MVV have focused on the 1514 virus. The initial approach which was selected was to make use of a transfection assay to determine the activity of various LTR forms. During the sequencing of the British MVV isolate, EV-1, differences in both LTR sequence and LTR organisation were observed between the two viruses EV-1 and 1514. Further, within the EV-1 population a number of LTR variants were observed. The activity of these sequences was examined in order to determine the possible role of variation in the LTR activity within the viral population. Comparison of the LTR sequences derived from the experimentally infected animals with those present within the infecting population would allow conclusions to be drawn on the degree of LTR selection *in vivo*. As the EV-1 variants included LTRs bearing duplications of various regions of the LTR it was also of interest to use these sequences to identify functional sequences within the LTR.

CHAPTER TWO

MATERIALS AND METHODS

2.1. Materials

Unless otherwise stated, the chemicals used for this thesis were obtained from Sigma Chemical Co., Poole, Dorset, UK. Restriction endonucleases and other DNA modifying enzymes, were purchased from Northumberland Biologicals Ltd., Cramlington, U.K., New England Biolabs., Hitchin, UK, Boehringer Mannheim, BCL, Lewes, UK, or Promega, Southampton U.K., unless stated otherwise. Radiochemicals were supplied by Amersham International PLC, Amersham, UK.

2.2 Bacterial Cultures

2.2.1 Bacterial Growth Media and Plates

The bacterial growth media used throughout this thesis was Luria Bertani (LB) broth. This media was supplemented with ampicillin (100µg/ml) for the selection of transformed cells, only plasmid vectors encoding the ampicillin resistance gene were used in these studies. LB/agar plates were made with LB/1% (w/v) agar (Oxoid, Basingstoke, England), LBamp plates were supplemented with 100µg/ml ampicillin.

Blue/white selection using the plasmid pCRII from the TA cloning kit (Invitrogen Corp., San Diego, California) was performed on LBamp plates with 200µg/ml 5-bromo-4-chloro-3-indolyl-ß-D-galactoside (X-gal). This vector was transformed into the supplied INVαF' competent E.coli which do not require IPTG (Isopropyl-ß-D-Thiogalactoside) to induce the expression of ß-gal.

2.2.2 Bacterial Strains, Stocks and Culture

For general subcloning, *E.coli* strain JM101 ([F' traD36 lacI^q lacz M15] proAB supEthi lac-proAB, Yannish-Perron et al 1985) was used for transformations. For vectors used in transfection experiments, plasmids were transformed into competent DH5α cells (Gibco BRL, Paisley, Scotland). Untransformed JM101 cells were cultured in LB and LBamp after transformation. For long term storage bacterial stocks were kept at -70^{o}C in 40%(v/v) glycerol.

2.2.3 Preparation and Transformation of Competent Bacteria

One ml of a saturated overnight culture was added to 40ml LB and grown until OD_{600} was between .35 and .40. For each transformation, 1.4ml of cells were spun down and resuspended in 0.5ml MR (10mM MOPS pH7.0, 1mM$RbCl_2$, 38mM$CaCl_2$). The cells were spun down again, resuspended in 0.5ml MRC (100mM MOPS pH6.5, 1mM$RbCl_2$, 38mM$CaCl_2$) and left on ice for 30 minutes. After this incubation, the cells were spun and resuspended in 0.15ml MRC and 3µl of DMSO was added. DNA was then gently mixed in and the cells were left on ice for 1hr. The tubes were heat shocked for 35 seconds at 55°C, and then returned to ice for 2 minutes. 1ml of LB was added to the cells and they were incubated at 37°C for 1hr and then plated on LBamp plates. The plates were checked 14-24hrs later for transformed colonies. Manufacturers competent cells were used for a number of procedures as described later in materials and methods.

2.3 Purification and Quantitation of Plasmid DNA

Small scale purification of plasmid DNA (minipreps) for screening transformed bacteria for the correct plasmid DNA was performed using an alkali lysis procedure (Birnboim & Doly 1979). 1.5ml of a 10ml overnight culture of *E.coli* was pelleted in a microfuge. The bacterial pellet was resuspended in Lysis buffer (25mM Tris pH8, 10mM EDTA, 50mM sucrose, 1mg/ml lysozyme). Two volumes of alkaline SDS (0.2N NaOH, 1% SDS) were added and mixed gently to avoid frothing until the solution was viscous. After the addition of 1.5 volumes of 3M NaAc (pH4.8) and the chromosomal DNA, SDS/protein precipitate and high molecular weight RNA was spun out by centrifugation at 12,000xg for 20 minutes in a benchtop microfuge. Nucleic acid in the supernatant was precipitated by addition of 0.6 volumes isopropanol. Precipitated DNA was collected by spinning at 12,000g for 15 minutes and resuspended in water or TE. The DNA was extracted with phenol:chloroform:isoamyl alcohol (P:C:I,49:49:2, pH7.5). After the first extraction, the RNase A was added (final concentration 100µg/ml) and incubated at 37°C for 30 minutes. After digestion of RNA the sample was extracted twice with P:C:I and once with Chloroform:Isoamyl Alcohol (49:1), the DNA was precipitated from the aqueous phase. Plasmid DNAs for use in transfection experiments were prepared using Qiagen Columns (Qiagen Ltd.) to ensure reproducibility and purity of plasmid preparations. For transfection plasmid DNA with A 280:260 of >1.8 was used.

The concentration of DNA in a sample was routinely determined by running a 10% aliquot on a 0.8-1.2% agarose gel. Minipreps gave a yield of 5-10μg. For transfections, DNA was also quantitated using a spectrophotometer. An absorbance at 260nm of 1.0 is equivalent to 50μg/ml double stranded DNA. The OD260/280 ratio gives a measure of sample purity, where a ratio of 1.8 is indicated a sample is free from protein or phenol contamination (Sambrook et al 1989).

2.4 DNA Sequencing

Sequencing was carried out using the dideoxy chain termination method of Sanger et al (1977). Double stranded DNA sequencing was performed after an alkali denaturation protocol and using the Sequenase kit (USB Corporation). Plasmid DNA (1-5μg) in a volume of 20μl was denatured by the addition of 2μl denaturing mix (2M NaOH, 2mM EDTA). After 5 minutes at room temperature the solution was neutralised (6.2μl 3M NaAc pH5.2) and DNA precipitated (7μl H_20 and 75μl 100% EtOH). Pelleted DNA was washed with 70% EtOH and resuspended in 7μl H_20, 2μl 5x Sequenase buffer and 1μl sequencing primer (0.1 pmol). Annealing was carried out by heating to 65°C for 2 minutes followed by cooling to about 35°C over 20-30 minutes. Full reaction volumes were collected by a brief spin. Sequencing reactions were performed using a USB Sequenase kit as per manufacturer's instructions. Briefly, 1μl 0.1M DTT, 2μl 1/5 Labelling mix, 0.5μl α^{35}S-dATP and 2μl 1/8 sequenase were added. The reaction was incubated at room temperature for 5 min. After extension, 3.5μl of the reaction was added to each of four tubes containing 2.5μl of Termination mix (A,C,G,T) and incubated at 37°C for 5 minutes. Reactions were terminated by the addition of 4μl of Stop mix. When compressions were observed in the DNA sequence reactions were repeated in the presence of 10% DMSO (labelling mix was diluted1/5 in 50% DMSO).

Biorad Sequi-Gen gel electrophoresis equipment was used for resolving the products of the sequencing reactions. Sequencing gels were 0.5xTBE (1x 89mM Tris-Borate, 89mM Boric Acid, 2mM EDTA), 6% Acrylamide (29:1 Acrylamide:Bis), 6M Urea. Gels were pre-run until temperature reached 55°C. Sequencing samples were heated to 75°C for 2 min before running and loaded at 2μl per well. Gels were run for between 2 to 8 hours to

resolve the products of the sequencing reactions. Prior to autoradiography, using X-OMAT AR film, gels were dried down on Whatman no.3 paper.

Sequence data was analysed using version 7.3 of the University of Wisconsin Genetics Computer group sequence analysis software, UWGCG 7.3 (Devereux et al 1984).

2.5 Restriction Digests

Restriction digests were performed in manufacturer's specified buffers. For double digests the buffer giving maximum activity for both enzymes was selected from manufacturer's tables. Generally 1-5μg of plasmid DNA was digested with 6-10 units of restriction enzyme for 2hrs at 37^{O}C.

2.6 DNA Ligation

2.6.1 Preparation of Vector

Vector DNA for use in ligation reactions was linearised by digestion with the appropriate restriction enzyme. In cases where a double digestion was performed single enzyme controls were included. After linearization, vector DNA was treated with calf intestinal phosphatase (CIP) to remove 5'Phosphate groups which greatly reduces the incidence of vector self ligation (Maniatis). The DNA was incubated with CIP (~.01unit per μg DNA) for 30 minutes at 37^{O}C.

2.6.2 Preparation of Inserts

Insert DNA was excised from the relevant plasmid by digestion with restriction enzymes and purified by separation through agarose gel. Two mechanisms of DNA purification from gels were used. For fragments over 600bp a Geneclean II kit was used. Where the band was below 600bp, the band containing the insert was excised from the gel and spun through glass wool, at 6500rpm for 15 min in a microcentrifuge. The insert was then precipitated and an aliquot used for quantification by electrophoresis.

When synthetic oligonucleotides were used as inserts they had first to be phosphorylated. This was performed by incubating 100pmol of oligonucleotides with ATP (1mM) and T4 polynucleotide kinase (10 units) for 10 minutes at 37^{O}C. For ligation, 50-100ng of dephosphorylated vector was used in the presence of a 10 fold molar excess of insert.

2.6.3 Ligation Reactions

Cohesive end ligations were carried out overnight at room temperature with a ligase concentration of 3-5 weiss units and manufacturers buffer. Blunt end ligations were carried out at 4^{o}C overnight using a ligase concentration of 6-10 weiss units. In general, 50-100ng of vector DNA was used for each ligation reaction with a three fold molar excess of insert. PCR reaction products were cloned using a TA cloning vector, pCRII (Invitrogen corp.), as per manufacturers instructions.

2.7 Agarose Gel Electrophoresis

Agarose gels of between 0.5 and 1.5% were used depending on the size of fragments being separated. Standard low IEEO agarose was used for all separations. A 6x loading buffer was used for all samples (0.25% bromophenol blue, 0.25% xylene cyanol and 15% Ficoll in H_2O). Gels were run in 1x TAE buffer (0.04M Tris-acetate, 1mM EDTA, pH7.6) containing 0.5µg/ml EtBr at between 50 and 80 volts.

2.8 PCR

PCR reactions to amplify LTR sequences were carried out in 100µl volumes in 0.5ml sterile eppendorf tubes. Reactions were performed in 1x PCR buffer (10mM Tris pH8.8, 1.5mM $MgCl_2$, 3mM DTT, 50mM KCl, 17µg/ml BSA, 1mM dNTPs) with ~50ng target DNA, 1unit *Thermus Aquaticus* (Taq) polymerase and primers (25-50pmoles). Cycling was performed using a Techne Programmable Driblock PHC-1. Amplification was performed over 35 cycles of the following conditions; 95^{o}C/0.6 minutes, 54^{o}C/0.5 minutes and 72^{o}C/1 minute. This was followed by 7.5 minutes at 72^{o}C to allow completion of partial extension products. All primers were supplied by OSWEL DNA Sevices (Dept. of Chemistry, Kings Buildings, West Mains road, Edinburgh). After amplification, 10% of the reaction product was analysed on an agarose gel as described in section 2.6.

2.9 Cloning into CAT Reporter Plasmids

2.9.1 CAT vectors

Two CAT reporter plasmids were used; pCAT12 (Thompson *et al.* 1994) which contains no promoter sequences in front of the Tn9 CAT gene and Δ56(a generous gift from Dr. J. Quinn) which contains 56 bases of the c-fos promoter. The vector pCAT12 was used to assay activity of the LTR clones (Section2.9.2) these contained complete promoter sequences, including a TATA box. The vector Δ56 was used to examine functional roles for the various LTR oligos.

2.9.2 Cloning Full Length LTRs into pCAT12

The following full length LTRs were cloned into pCAT12; EV-1 LTRs 1, 48, 44, 30, 28, 19, 18 and the 1514 LTR. EV-1 LTRs 30, 19, 18 and 28 were in the correct orientation for direct cloning from pTZ18R by double digest with Sac1/BamH1. After removal from pTZ18R, the LTRs were gel purified and ligated into pCAT12.

EV-1 LTRs 1,48 and 44 were in the wrong orientation for direct cloning so were cloned by PCR using primers 974S and 975S. These primers were designed to match those used in the original LTR cloning with the addition of restriction enzyme sites. There was no existing 1514 LTR clone so a PCR was carried out using a HIRT preparation of 1514 infected cells (provided by Dr. D. Roy) and primers 974S and B261. The primers used for cloning the LTRs are shown in Fig.2.1.

Figure 2.1 Primers for Amplification and Sequencing of LTRs

974S	CCCGAGCTCACTGTCAGGGCAGAGAAC	(Sac1)
975S	CCCGGATCCACAGCCAACTCCTTTATTAAC	(BamH1)
B261	CCCGGATCCCTTTATTGAGCTTTTCAGGCA	(BamH1)
309X	GATGCTTAAGTCATAACC	
310L	CATGATTGAGAATGAC	

Amplification products were cloned into pCRII (Invitrogen TA cloning kit) and sequenced. The pCRII vector contains annealing sites for M13 sequencing primers. These were used to confirm the sequence of the cloned DNA. After confirmation that the PCR

cloned LTRs were of the correct sequence and the absence of errors introduced by PCR, these inserts were cut out and inserted into pCAT12. The orientation of LTRs for all pCAT vectors was confirmed using an internal primer 310L.

2.9.2 Cloning of partial LTR sequences

Two EV-1 LTRs, LTR-1 and LTR-48, were used to generate truncated LTR CAT vectors. PCR was performed using the full length LTR as target and primers 975S and 309X (Fig.2.1). The amplified products were cloned into pCRII, sequenced, then transferred to pCAT12.

2.9.3 Oligonucleotides and Cloning into Δ56

Double stranded oligos were cloned into the vector Δ56. Oligonucleotides were inserted upstream of the c-fos TATA box sequence using a unique Sal I site in the vector. A list of the oligonucleotides used in these experiments are outlined in Fig.2.2. All of these oligonucleotides, with the exception of the 1514 AP-1 and EV-1 AP-1 sequences, have Sal I compatible overhangs to facilitate cloning into Δ56.

Ligations were carried out using phosphorylated oligonucleotides (Section 2.5). Transformed colonies were screened for insert by digestion with Sal I. Insertion of oligonucleotide into the plasmid results in the loss of the Sal I site, allowing screening of samples prior to sequencing. Cloning efficiency varied from 30-80%. All colonies where insert was present were sequenced across the Sal I site. These vectors are shown in Fig.2.3.

2.10 Cell Lines and Tissue Culture

The primary chondrocyte cell line used in this thesis (supplied by Lawrence Dickson) was derived from the carpal joint of a MVV negative sheep and grown in culture as fibroblast-like monolayers. The cells were grown to pass 5 and then frozen to supply a bank of cells for future experiments. For storage, the cells were frozen as 1ml aliquots in freezing mix (91% serum 9% DMSO). The cells were slowly frozen to -70°C overnight then placed in liquid nitrogen for long term storage. All experiments described in this thesis used these cells at between passes 10 and 20. Cells were cultured in DMEM supplemented with 10% FCS, penicillin (100 units/ml), streptomycin (100μg/ml) and L-glutamine (2mM). Cells were grown until monolayers reached confluence then split at either 1:3 or 1:4. Cells were trypsinized with

FIGURE 2.2 Synthetic oligonucleotides

Control AP-1	TTCCGGCTGACTCATCAAGCG	(B260)
	AAGGCCGACTGAGTAGTTCGC	(B619)
1514 AP-1	TGCTTGAGTCATAACC	(B533)
	ACGAACTCAGTATTGG	(B532)
EV-1 AP-1	TGCTTAAGTCATAACC	(C146)
	ACGAATTCAGTATTGG	(C145)
1514 oligo	TGATGCTTGAGTCATAACCGCA	(B383)
	ACTACGAACTCAGTATTGGCGT	(B384)
EV-1 *ex vivo*	TGATGCTTAAGTCATAACCGCA	(B535)
	ACTACGAATTCAGTATTGGCGT	(B534)
EV-1 *ex vitro*	TGATGCTTAAGTCATAACCACA	(B245)
	ACTACGAATTCAGTATTGGTGT	(B244)
VSS oligo	GAGTCATAACCGCA	(C144)
	CTCAGTATTGGCGA	(C143)
Upstream Region (USR)	TCAGGATGACACAGCAAATGTAACCGCAAGTTCTGCTT	(E128)
	AGTCCTACTGTGTCGTTTACATTGGCGTTCAAGACGAA	(E129)
Tα4	CCCCCCAACCGCAGGTGCAG	(M0789)
	GGGGGGTTGGCGTCCACGTC	(M0790)

The synthetic oligonucleotides used in this work are listed above. With the exception of the 1514 and EV-1 AP-1 sequences all possess Sal 1 overhangs (TCGA) for cloning into the Δ56 vector. The oligonucleotides are identified by their Oswel codes. The positioning of these sequences is discussed fully in Chapter 4.

FIGURE 2.3

The Δ56 vectors generated during this work are listed. The orientation is given relative to the *c-fos* TATA box. Arrows denote orientation of oligonucleotides with relation to the basal promoter. Each arrow represents a single copy of the oligonucleotide.

Δ56 VECTORS

1514 full length oligo

clone no.	orientation
1	→
2	←
8	→
9	→
16	←
17	→
20	←

EV-1 ex vivo

clone no.	orientation
4	→
6	→
9	←
15	→
18	←

EV-1 ex vitro

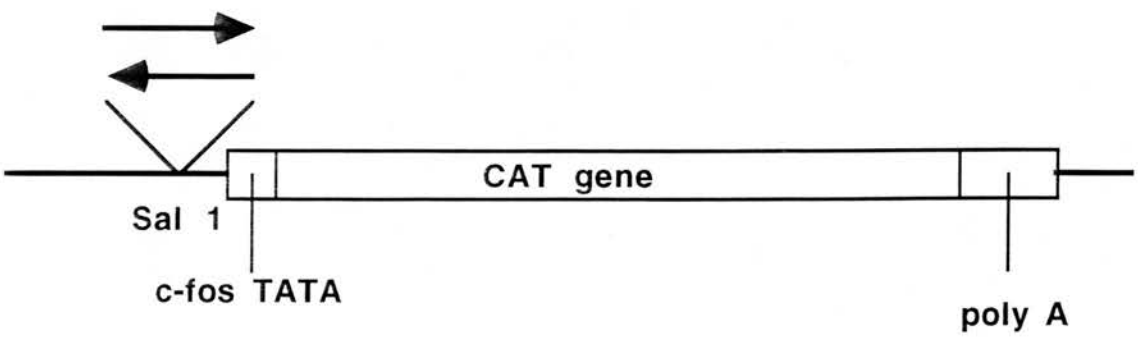
clone no.	orientation
1	→
3	→
4	→
5	←
8	←

VSS oligo

clone no.	orientation
1	↔
2	→ ←
4	→ ←
5	←
6	← ←

1514 Upstream Region

clone no.	orientation
4	→



Trypsin/Versene (0.05% w/v trypsin: 0.02% w/v versene) after having first been washed with 0.02% versene to remove all culture media. For transfection, chondrocytes were split into 60mm dishes and grown to 80-90% confluence.

HeLa Ohio cells were obtained from the European Collection of Animal Cell Cultures (ECACC, Porton Down, Salisbury). These cells were cultured in DMEM supplimented with 5% FCS, penicillin (100 units/ml), streptomycin (100µg/ml) and L-glutamine (2mM). Cultures were grown to confluence then split 1:10 as described above. Stocks of HeLa cells were frozen in liquid nitrogen using the same protocol as described for the chondrocytes.

2.11 Transfection and Reporter Gene Analysis

2.11.1 Transfection

Transfection was performed using a calcium phosphate precipitation protocol. Immediately prior to transfection, media were aspirated and replaced with fresh media containing 5% FCS. Each 60mm dish (~3 x 10^6 cells) was treated with 0.5ml DNA/calcium phosphate precipitate (5µg plasmid/plate). Each 1ml of precipitate was composed of 500µl 2 x DNA precipitation buffer (50mM Hepes pH7.05, 1.5mM Na_2HPO_4, 10mM KCl, 280mM NaCl, 12mM Glucose), 62µl 2M $CaCl_2$, 10-20µg DNA and dH_2O to a final volume of 1ml. The DNA and dH_2O and $CaCl_2$ were placed in a sterile tube and the 2x precipitation buffer was added slowly with constant gentle mixing. The samples were then left to sit at room temperature for 25-30 minutes to allow the precipitate to form. The precipitate was then added to the culture medium of each plate. The plates were incubated for 4 hours.

After the incubation, the media were aspirated from the plates and the monolayers washed with 4ml prewarmed PBS. The cells were then shocked with 15% glycerol (in 1x precipitation buffer) for 1 minute. The shocking buffer was removed and replaced with culture medium. Cells were harvested 44 hrs post transfection.

2.11.2 Preparation of Cell Extracts for Enzyme Assays

Whole cell extracts were prepared from transfected cells by a freeze thaw method. Cells were scraped from plates after washing in PBS containing 1mM EDTA. Harvested cells were kept on ice. Cells from each plate were collected in 1.5ml and spun down, the pelleted cells were then resuspended in 250mM Tris pH8.0.

Lysis was performed by 3 cycles of dry ice/methanol to 37°C water bath. Debris was removed by a 5 min 13500 rpm microcentrifuge spin. CAT and ß-Gal activities in this extract were then measured.

2.11.3 Assay for CAT Activity in Cell Extracts

CAT activity in the extract was determined by liquid scintillation counting as described by Seed & Sheen (1988). Briefly,10µl of extract was mixed with 5µl n-butyryl CoA (5mg/ml), 2µl ^{14}C-chloramphenicol (0.025mCi/ml) and 113µl 250mM Tris pH8.0. After incubating for 3hrs at 37°C the reactions were placed on ice and extracted with 250µl mixed xylenes (Aldrich Chemical co., Dorset England). The tubes were mixed well then spun at 13500rpm in a microcentrifuge for 3 minutes. The upper organic phase was removed and extracted with 250µl H_20. After this back extraction 125µl of the organic phase was placed in a scintillation phial, mixed with scintillation fluid and counted for radioactivity.

2.11.4 Assay for ß-Gal Activity in Cell Extracts

The reporter plasmid pSVßgal (Promega) was used as a control for transfection efficiency. This plasmid contains the ß-gal gene under the control of the SV-40 promoter and enhancer, to give high level expression. Extracts were assayed for ß-Gal activity as follows. 50µl of extract was mixed with 50µl of 2x ß-Gal assay buffer (120mM Na_2HPO_4, 80mM NaH_2PO_4, 2mM $MgCl_2$, 100mM ß-mercaptoethanol, 1.33mg/ml ONPG [o-nitrophenyl-ß-D-galactopyranoside]) in a flat bottom microtitre plate (Dynatech corp.). The plate was incubated at 37°C for 1-2 hrs or until a strong positive reaction was observed. Reactions were stopped by the addition of 150µl 1M $NaCO_3$. ß-Gal activity was measured by determing OD_{405} on a microtitre plate reader (Dynatech corp.).

2.12 Preparation of Nuclear Extracts

Nuclear extracts were prepared as described by (Schreiber *et al.* 1989).This is a modified version of the protocol described by Dingman *et al.* (1983). Briefly, 1x10^7 cells were collected and resuspended on ice in 400µl ice cold buffer A (10mM Hepes pH7.9, 10mM KCl, 0.1mM EDTA, 0.1mM EGTA, 1mM DTT). After a 15 minute incubation in this hypotonic buffer,

the cells were lysed by the addition of 25µl of 10% Triton X-100. The sample was vortexed for 10 seconds and the nuclei were then immediately spun out and resuspended in 50µl of buffer B (20mM Hepes pH7.9, 0.4M NaCl, 1mM EDTA, 1mM EGTA, 1mM DTT).

The nuclei were then rocked vigourously at 4°C for 15 minutes. Finally the nuclei were spun out and the nuclear extract was aliquoted and stored at -70°C. Prior to freezing, protein concentration was determined (Biorad protein assay kit) and the extract diluted to 5µg/µl using buffer B. All buffers for nuclear extract contained the following proteolytic inhibitors; 5µg/ml aprotonoin and leupeptin, 2µg/ml pepstatin and 1mM PMSF (Phenylmethylsulphonyl fluoride), from a fresh 50mM stock in isopropanol.

2.13 Gel Retardation

Gel retardation assays were performed essentially as described by Goodwin (1990) with minor modifications.

2.13.1 Generation of Radiolabelled Probes

Target DNA sequences for use in the assay were radiolabelled using T4 polynucleotide kinase and γ^{32}P-ATP. Unincorporated γ^{32}P-ATP was removed by spin column chromatography through Sephadex G25. Approximately 50,000 cpm of probe were used per reaction.

2.13.2 Gel Retardation Assay

Reactions were performed as follows. In a microcentrifuge tube, 1µl each of nuclear extract (5µg), poly [dI-dC.dI-dC] (5µg) and 5x retardation buffer (20% glycerol, 50mM Hepes pH7.6, 5mM MgCl2, 2.5mM EDTA, 2.5mM DTT) were mixed. For all reactions, the final concentrations of NaCl and Glycerol were 80mM and 4% respectively. The competitor poly [dI-dC.dI-dC] was used to reduce background generated by non-specific protein-DNA interactions. The poly [dI-dC.dI-dC] was sonicated and annealed (95°C then slow cooling) to provide material with a large number of free ends. The size of this material was checked by agar gel electrophoresis (50-500b.p). Reactions were left for a 5 minute preincubation to allow the removal of non-specific DNA binding proteins. Specific binding proteins were then removed by the addition of competitor oligonucleotides. After a further 5 minute incubation,

1µl of radiolabelled probe DNA was added to the side of each tube. Mixing was performed by a brief microcentrifuge spin. Samples were left for a further 10 minutes. All incubations were carried out on ice.

Complexes were separated by electrophoresis through a 4% non-denaturing polyacrylamide gel in 0.4 x TBE buffer. An acrylamide:bis ratio of 29:1 was used, and gels were cast using 1.5mm spacers. Electrophoresis was carried out at 4^{o}C with recirculation of the buffer. The gels were run at a constant 150V and prerun for 30 minutes prior to addition of the samples. At this stage, 5µl of tracking dye was added to the wells in order to monitor the running of the gel. To limit diffusion and dissociation of complexes, the samples were applied to the gel with the current flowing. Gels were run for ~3hrs or until the bromophenol blue band was ~3cm from the bottom of the gel. After electrophoresis the gels were dried on Whatman no.3 paper. Dried gels were exposed to Kodak X-OMAT AR film for 1hr to 3 days depending on the strength of the signal.

For gel super shift experiments, complexes were formed as described above. After incubation with radiolabelled probe, antibody was added and samples incubated on ice for 30 minutes. Samples were then applied to the gel as described above. The rabbit anti-jun antibody was obtained from Santa-Cruz Biotech. (San Diego, USA). This antibody was generated against a peptide from murine c-Jun and is reactive against all Jun family members. A control rabbit IgG was obtained from Sigma.

2.14 UV Crosslinking

UV crosslinking of DNA to protein was performed essentially as described by Ausubel *et al.* (1993). The target oligonucleotide was BrdU substituted and hybridised to the corresponding sequence to generate a double stranded oligonucleotide (Fig.2.4).

Fig.2.4 Probe for UV crosslinking

Oligo		
C144	TCGACAGTCATAACCGCA	
HOO75	GTCAGTAXTGGCGXAGCT	X= 5 Bromo dU

2.14.1 Formation of DNA-Protein Complexes

Complexes were formed as in the standard gel retardation protocol (Section 2.13.2). With the exception of using the BrdU substituted oligo as probe. Crosslinking was carried out using a UV Stratalinker 1800 (Stratagene Ltd.). Samples were placed, on ice, at a standard distance from the UV tubes. Sample exposure to UV was timed using the Stratalinker's built in timer, this allowed consistent exposure between experiments.

Aliquots were removed at various time points and mixed with SDS reducing buffer (150mM ß-mercaptoethanol, 15% glycerol, 0.05% SDS, 0.005% (w/v) bromophenol blue, 50mM TrisHCl pH6.8). Reactions were assembled to give 5µg of protein and 50,000cpm per aliquot. Optimisation of crosslinking conditions is described in Chapter 4.6.2.

2.14.2 SDS-PAGE of Crosslinked Complexes

Complexes were resolved by discontinuous polyacrylamide gel electrophoresis run under reducing conditions (Laemmli 1970) using Biorad Mini-protean II vertical slab gel apparatus.

A stock solution of 30% acrylamide:0.8% bisacrylamide was diluted to 10% in 250mM TrisHCl pH8.7, 1.3mM EDTA and 0.1% SDS for the main gel. Polymerisation was carried out by adding 0.5% APS (ammonium persulfate) and 0.05% TEMED. Gels were poured using 0.75mm spacers. The stacking gel contained 3.5% stock acrylamide, 145mM TrisHCl pH6.8 and 0.1% SDS. After polymerisation, the samples were electrophoresed at 200V in 25mM TrisHCl, 6mM EDTA, 0.1% SDS and 156mM glycine. All samples including C^{14} labelled markers (Amersham) were boiled for 3 minutes in SDS reducing buffer immediately prior to electrophoresis. After electrophoresis, the gels were treated with EN^3HANCE (NEN research products, Boston, USA) to improve detection of C^{14} labelled MWT markers. After this treatment, the gels were dried onto Whatman no.3 paper and autoradiographed at -70°C using Kodak X-OMAT AR film.

CHAPTER THREE

3.1 Introduction

As discussed in Chapter 1 the principal sequences governing transcription are positioned within the U3/R region of the LTR. Unlike HIV, efficient transcription in MVV does not appear to be dependent on the presence of the Tat protein. The importance of the trans-activator Tat in HIV is shown by the absolute dependency of viral replication on the presence of functional Tat protein (Dayton *et al.*, 1986, Fisher *et al.*, 1986). This absolute dependence is not seen in MVV.

While LTR variation has been described in MVV there has been no examination to date on the effects of this variation on LTR function. The significance of this sequence variation for viral replication and the infectious process *in vivo* is unclear in MVV. When LTR variation is followed after experimental MVV infection there is some limited evidence for selection of certain LTR sequences from the infecting population (Sargan *et al.*, 1995). The question then becomes: does sequence variation give rise to functional diversity in the LTR population? One important distinction between the ruminant lentiviruses, MVV and CAEV, and HIV is the distinct regulatory role of the Tat protein in the Ovine and Caprine viruses. There is still some controversy regarding the importance of Tat in these viruses and whether or not it is absolutely required for viral growth *in vitro* and *in vivo* (Saltarelli *et al.*, 1993, A. Harmache pers. com).

The MVV Tat protein is considerably less active than HIV Tat. While HIV Tat shows transactivation from a hundred to several thousand fold, depending on the cell line used (reviewed by Antoni et al 1994). The MVV Tat protein shows a less marked effect (Gourdou *et al.*, 1989, Gdovin & Clements 1992, Neuveut *et al.*, 1993). This difference may in part reflect different U3 promoter activities between the two viruses (Chapter 1). Estimations of MVV Tat transactivation vary from 5-40 fold, making the MVV protein between 10 and 100 fold less active than HIV Tat. This marked difference in transactivation potential indicates functional differences between the two proteins. In MVV the Tat target sequences appear to be located solely within the U3 promoter region of the LTR (Hess *et al.*, 1989) and there is no evidence for a TAR region (Hess *et al.*, 1989). Thus the sole targets for MVV Tat action appear to be the cellular factors which interact with the U3 region of the LTR. There is however no evidence to date for either direct or indirect binding of the MVV Tat protein to MVV LTR sequences (Gdovin & Clements 1992 & J. Clements pers. communication).

Variation in sites for transcription factors within the LTR does occur between the MVV isolates sequenced to date (Querat *et al.*, 1990, Sargan *et al.*, 1991, Andresson *et al.*, 1992). As the U3 LTR region appears to represent the sole target for the regulation of MVV transcription variation within the U3 promoter region of the LTR may thus be of greater significance than in HIV.

As a number of LTR variants had been identified during the sequencing of EV-1 these LTRs were used to test the effects of sequence variation on the rate of transcription from the LTR. These experiments were performed by inserting LTR sequences upstream of the CAT reporter gene in a transfection vector.

3.2 Overview of EV-1 LTR Variants

3.2.1 LTR Populations in vitro and in vivo

The MVV biological clone, EV-1, was derived from explant culture of a sheep showing Maedi disease and after several passes in tissue culture this virus was sequenced (Sargan *et al.*, 1991). This initial isolate was also used to generate laboratory stocks of virus for future experimental infections. The population of LTRs derived from this *in vitro* virus are shown in Fig.3.1 (from Sargan *et al.*, 1995). These LTR sequences vary due to point mutations but also show variable repeat structures. These repeats can be classified according to the region repeated and the repeat boundaries. This gives four basic repeat types (as illustrated in Fig.3.1). Type I repeats cover the upstream region of the LTR, and, depending on their size are classified as type IA or IB. The type II repeats cover an extended region of the LTR including AP-1/AP-4/AP-1 element. Type III repeats cover only this AP-1/AP-4/AP-1 region. In addition to this subdivision based on repeat structures the LTRs can also be subdivided via the presence or absence of an insert at position 90. This division is important when these *ex vitro* sequences are compared to those found *in vivo*.

Since the virus from which the initial LTRs were isolated was used for subsequent experimetal infections it was also possible to obtain LTR sequences from the lymph nodes and peripheral blood of two experimentally infected sheep (Sargan *et al.*, 1995). These *ex vivo* LTR sequences are shown in Fig.3.2. As the LTRs isolated from the infected animals were derived from the tissue culture virus, these two populations are directly related. The relationship between these two LTR populations can be seen in two ways. Either the *ex vivo*

FIGURE 3.1 LTR variants obtained from tissue culture

LTR sequences obtained following PCR from DNA of *in vitro* EV-1 infected cells are shown. Sequences are aligned against LTR 1 as this was the first sequence obtained. The four repeat types are given at the top of the figure along with the nucleotide positions.

The numbering system used throughout this thesis for the LTR sequences is as follows; position 1 corresponds to position 8955 of the published EV-1 sequence (Sargan *et al.*, 1991) and 8773 of 1514 (Sonigo *et al.*, 1985).

The sequence data presented in this figure and fig.3.2. was the work of I. Bennet.

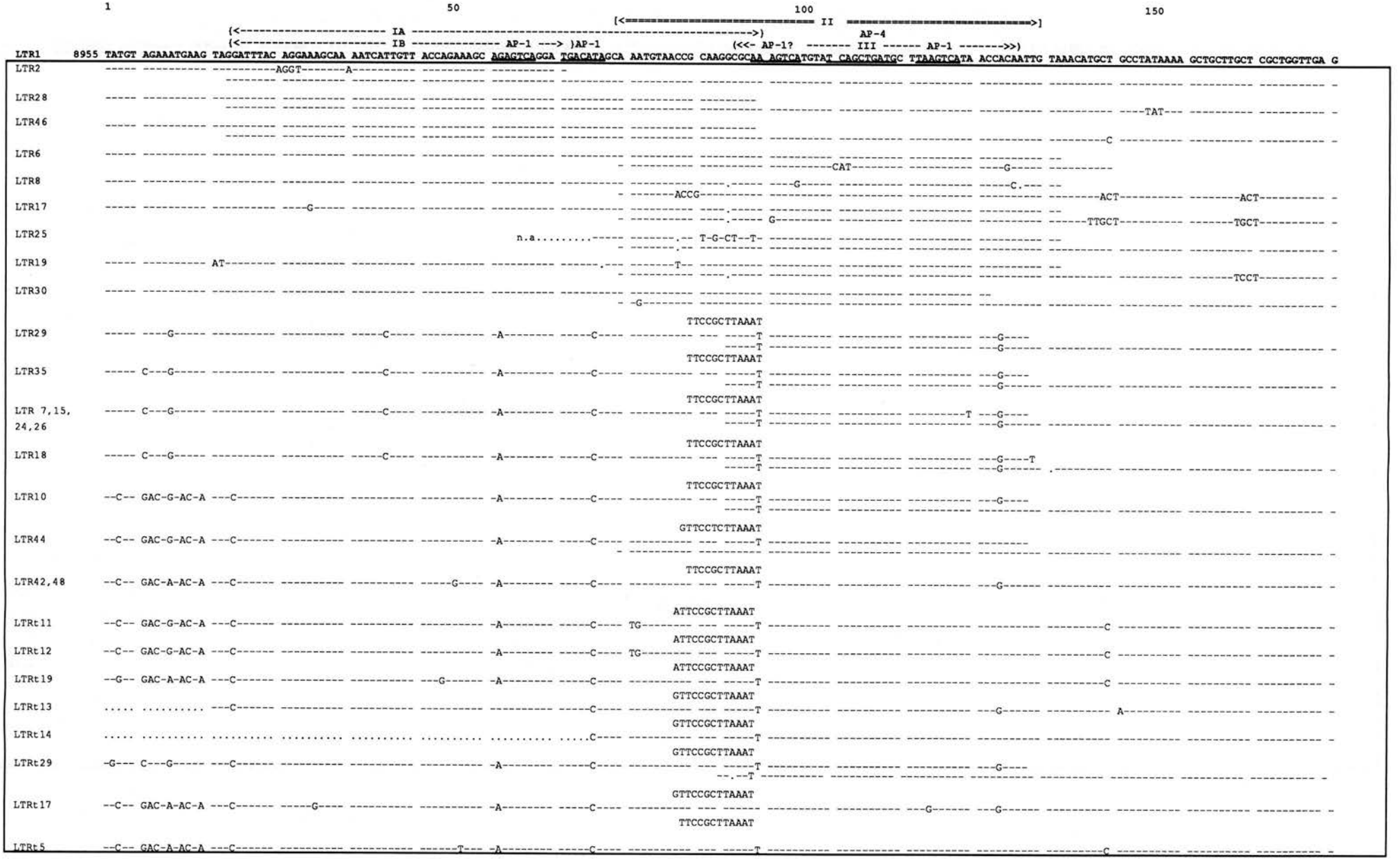
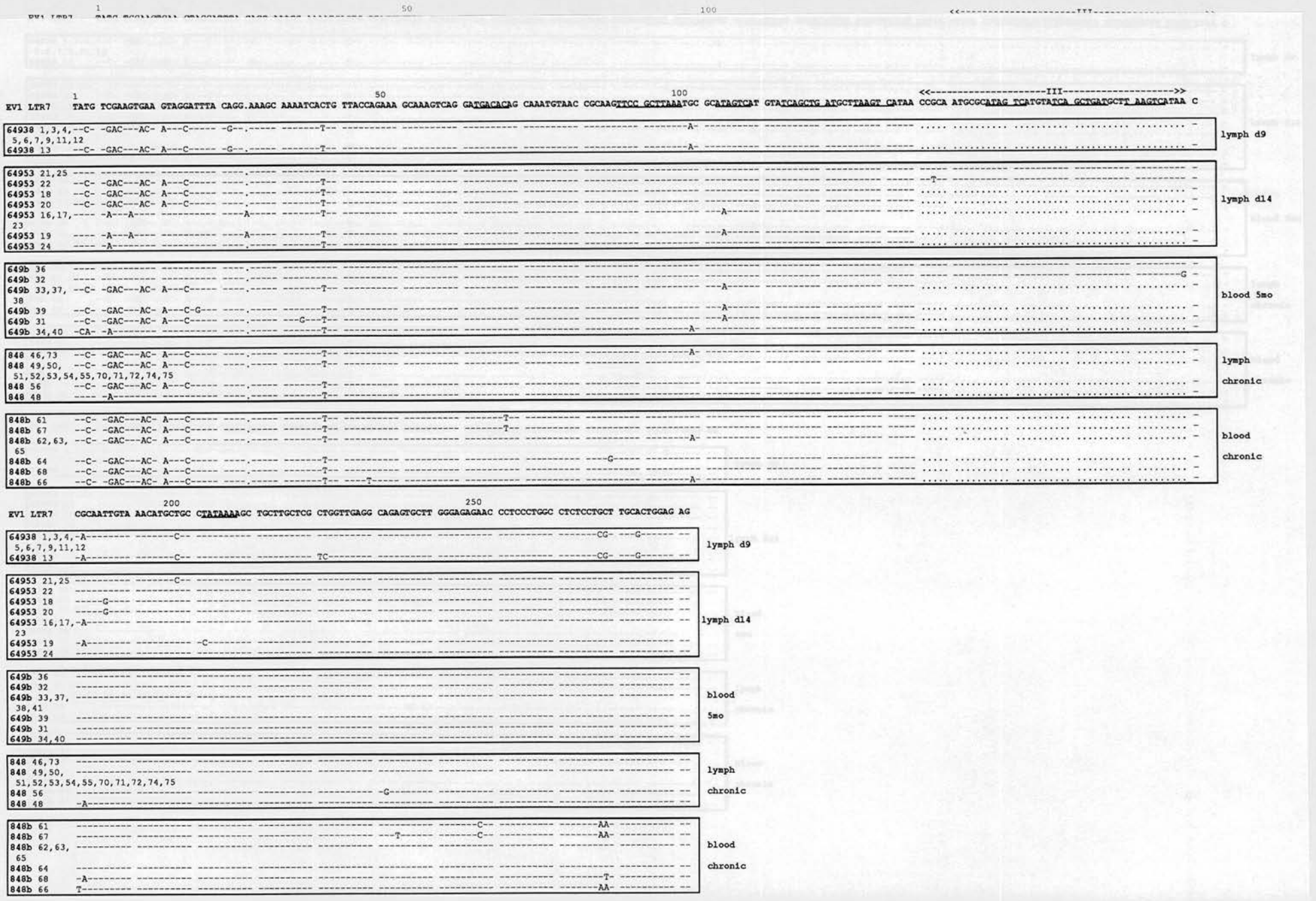

1
50
100
150
(<---------- IA ----------)
(<---------- IB ---------- AP-1 --->)AP-1
[<========== II ==========>]
(<<- AP-1? ------- III ------ AP-1 ------->>)
AP-4
LTR1 8955 TATGT AGAAATGAAG TAGGATTTAC AGGAAAGCAA AATCATTGTT ACCAGAAAGC AGAGTCAGGA TGACATAGCA AATGTAACCG CAAGGCGCAA AGTCATGTAT CAGCTGATGC TTAAGTCATA ACCACAATTG TAAACATGCT GCCTATAAAA GCTGCTTGCT CGCTGGTTGA G
LTR2
LTR28
LTR46
LTR6
LTR8
LTR17
LTR25
LTR19
LTR30
LTR29
LTR35
LTR 7,15,
24,26
LTR18
LTR10
LTR44
LTR42,48
LTRt11
LTRt12
LTRt19
LTRt13
LTRt14
LTRt29
LTRt17
LTRt5

FIGURE 3.2 LTR variants from experimentally infected sheep

LTR sequences obtained from two experimentally infected sheep are shown. Sheep 649 (top 3 panels) was infected and then used to obtain samples for PCR at defined time points after infetction, shown on the right hand side of each panel. Sheep 848, a chronically infected animal, was used to obtain data on LTR sequences late in infection. Repeat types, from Fig.3.1 are shown at the top of the figure along with nucleotide positions.


EV1 LTR7 TATG TCGAAGTGAA GTAGGATTTA CAGG.AAAGC AAAATCACTG TTACCAGAAA GCAAAGTCAG GATGACACAG CAAATGTAAC CGCAAGTTCC GCTTAAATGC GCATAGTCAT GTATCAGCTG ATGCTTAAGT CATAA CCGCA ATGCGCATAG TCATGTATCA GCTGATGCTT AAGTCATAA C
64938 1,3,4,5,6,7,9,11,12
64938 13
lymph d9
64953 21,25
64953 22
64953 18
64953 20
64953 16,17,23
64953 19
64953 24
lymph d14
649b 36
649b 32
649b 33,37,38
649b 39
649b 31
649b 34,40
blood 5mo
848 46,73
848 49,50,51,52,53,54,55,70,71,72,74,75
848 56
848 48
lymph
chronic
848b 61
848b 67
848b 62,63,65
848b 64
848b 68
848b 66
blood
chronic
EV1 LTR7 CGCAATTGTA AACATGCTGC CTATAAAAGC TGCTTGCTCG CTGGTTGAGG CAGAGTGCTT GGGAGAGAAC CCTCCCTGGC CTCTCCTGCT TGCACTGGAG AG
64938 1,3,4,5,6,7,9,11,12
64938 13
lymph d9
64953 21,25
64953 22
64953 18
64953 20
64953 16,17,23
64953 19
64953 24
lymph d14
649b 36
649b 32
649b 33,37,38,41
649b 39
649b 31
649b 34,40
blood
5mo
848 46,73
848 49,50,51,52,53,54,55,70,71,72,74,75
848 56
848 48
lymph
chronic
848b 61
848b 67
848b 62,63,65
848b 64
848b 68
848b 66
blood
chronic

LTRs are a subpopulation of the *in vitro* LTR sequences or they are rapidly evolved from the tissue culture population. The *in vivo* LTRs all have an insert present at position 88 suggesting that they are derived from the second, insert bearing, group of *in vitro* LTRs.

The variation between the *in vitro* and *in vivo* LTR populations does suggest a degree of selective pressure being exerted on LTR sequences within the infected animal since the *ex vivo* LTRs are all derived from the *in vitro* population. This idea of LTR selection during infection is dependent on both the *ex vivo* and *ex vitro* populations described here truly representing the actual LTR variants present within each milieu.

These LTR sequences cannot be taken as representing the entire population of those present in the population as only a single primer pair were used to generate the sequences shown in Fig.3.1. While the sequences described are primer dependent it is unlikely a given LTR structure or sequence is linked to the upstream and down stream primer sites given the high recombination rates seen in lentiviruses. This is discussed more fully in section 3.2.3.

3.2.2 PCR error and LTR variation

With such a large variation in LTR structure and sequence observed it is important to assess the possible contribution of PCR artifacts. Any cloning strategy based around PCR is error prone due to the large number of polymerisation cycles. The most common form of PCR artifact is base substitution, with purine to purine (A-G) transition being the most frequent. Other commonly encountered artifacts are single base deletions/insertions and inversions. Estimations as to the frequency of errors with PCR vary from 0.126% to 0.03% of bases (Meyerhans *et al.*, 1989, Perrin *et al.*, 1990, Carey & Dalziel 1994) depending on template and reaction conditions. The error rate for the conditions used to clone the MVV LTRs was calculated at 0.07% (1:1500) (Sargan *et al.*, 1994). Taking this error rate then the frequency of point mutations can be taken as just over one per 5 LTR sequences, the cloned sequence of these LTRs being between 300-350 base pairs. The significance of point mutations is hard to discern due to the problems posed by separating genuine variation from PCR artifact. The possible significance of these changes will be discussed separately in each context and will mainly feature in discussions of possible transcription factor binding sites where a single base change may have a sizable effect on site affinity for a particular factor.

While a number of the single base variations between these LTR sequences may be attributable to polymerase infidelity, the insertion and deletion of repeat structures or the addition/removal of larger inserts is less readily explained. Possible mechanisms for the insertion of repeats during PCR are outlined in Fig.3.3. This type of artifact would be dependent not on polymerase error but rather the mechanics of the PCR process and the nature of the target sequence, ie. the presence of small repeated motifs.

Loss of repeats from a sequence can occur via a reverse process to that outlined in Fig.3.3. In this case a sequence containing repeats is partially extended then dissociated, the first repeat can then anneal to the second allowing the generation of a truncated, repeatless, sequence. It should be noted that loss of repeats can be mediated by any sequence within the repeated region thus making repeat loss a far more efficient reaction than repeat insertion.

During the original cloning of the EV-1 LTRs this generation/loss of repeats was examined. PCR reactions were carried out using a single, cloned, LTR sequence as the template. LTRs which did or did not posses repeats were analyzed in separate reactions. Analysis of the products revealed that while it was possible to lose repeats during the amplification procedure, there was no evidence for their generation (Sargan *et al.*, 1995). It can be concluded with reasonable certainty that the presence of repeat structures within the two populations of virus is a genuine phenomenon.

3.2.3 Viral mechanisms for the generation of repeat structures

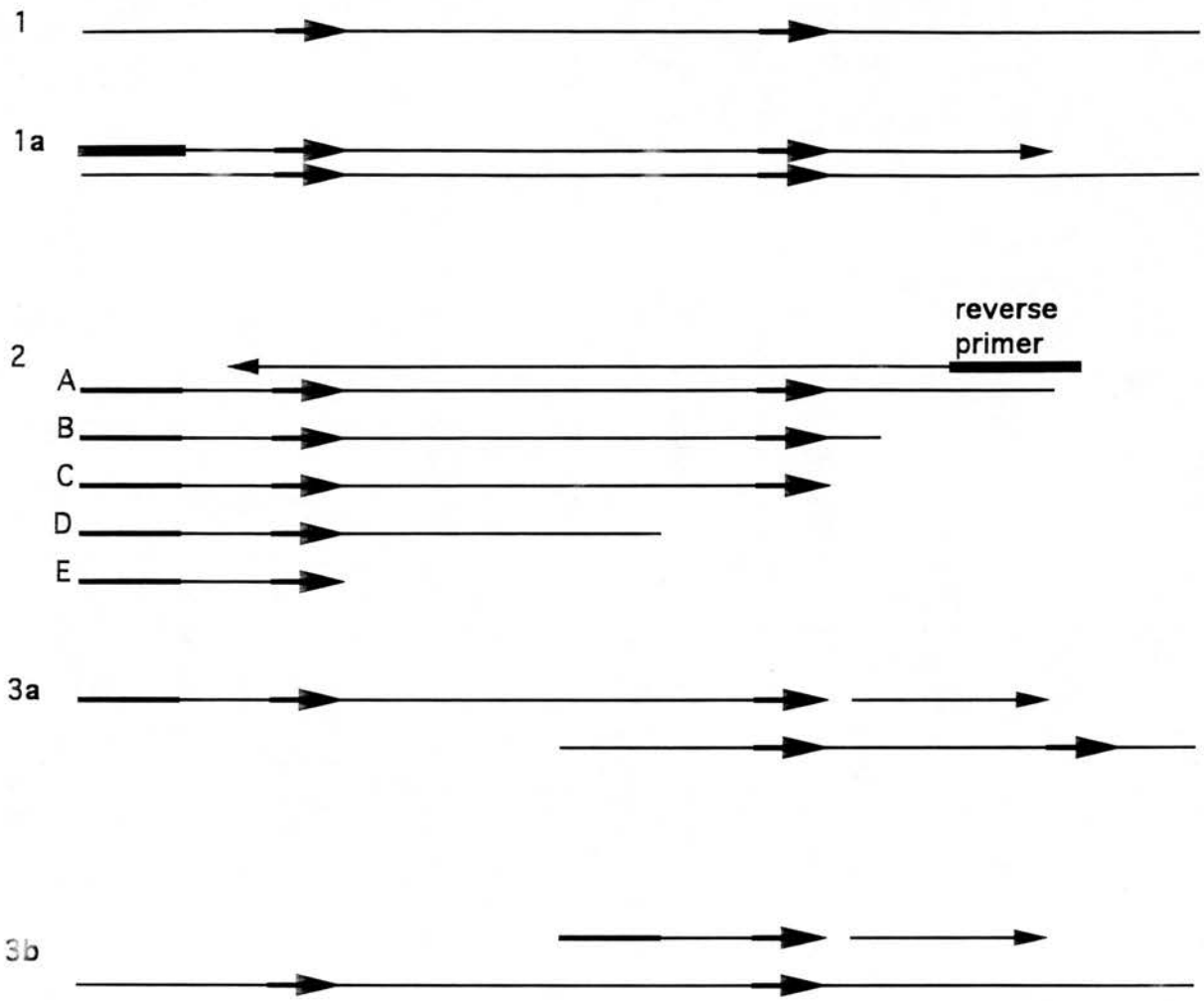
As the LTR variants observed appear to be products of the viral life cycle, both *in vivo* and *in vitro,* the question of how these structures arise must be addressed (assuming that a single progenitor LTR type is capable of giving rise to all the repeat structures discussed in these studies).

As PCR artifacts do not appear to be responsible for the generation of the repeat structures observed they must arise spontaneously during the MVV life cycle. One method for the insertion of repeat structures into nucleic acid is recombination between related sequences. Recombination is known to occur between distinct viral quasispecies in HIV infected individuals (Howell *et al.*, 1991). Recombination can theoretically occur at two stages of the retroviral lifecycle - during reverse transcription or between two proviral DNA sequences after superinfection. When these two possible mechanisms are tested experimentally only co-

FIGURE 3.3 Schematic outline for the introduction of repeats via PCR

A template containing repeated sequences (1) is extended from a 5' primer. This gives rise to a number of extended species (2). Only those extended beyond the site for the reverse primer (2A) act as templates for the second, reverse primer.

In the case of partially extended products which terminate at the end of the first (2E) or second (2C) repeat these can then reanneal to the full length template. This will give rise to species with a repeat (3a) or deletion of the sequence between the two repeats (3b).


1
1a
reverse
primer
2
A
B
C
D
E
3a
3b

infection leads to the generation of recombinant viruses (reviewed in Coffin 1979).

Recombination during reverse transcription is possible due to the presence of two viral genomes and the ability of the reverse transcriptase to jump when it reaches the end of a template, as it has to do when it reaches the end of one genome and switches to the other, or when it encounters a break in the strand it is processing (Junghans *et al.*, 1982, Coffin 1979). Reverse transcriptase normally transfers at the end of the genome via paired homologous sequences, however the enzyme can make strand transfers in the absence of extended homology. This step may be triggered by pausing of the polymerase (Buiser *et al.*, 1993). In this scenario misalignment of the two genomes would allow the introduction of repeats into the MVV sequence.

In the *ex vivo* sequences, only a single repeat type was found. Comparison of the boundary sequences in these LTRs shows a conserved sequence at the start and end of each repeat (TAACCGCAA) Fig.3.2. This sequence represents a possible region for template misalignment prior to recombination events. In the case of the repeats unique to the *in vitro* population, there is no evidence for conserved regions at the repeat boundaries. This may be due to a slightly modified system operating *in vitro.* Differences between the *in vitro* and *in vivo* environment will no doubt arise in relation to rate of viral replication within cells and multiplicity of infection (Haase *et al.*, 1982 , Harris *et al.*, 1984, Peluso *et al.*, 1985).

In vivo there is likely to be a very low multiplicity of infection with limited viral replication rates (Peluso *et al.*, 1985). In contrast, *in vitro* the high replication rates together with the high multiplicity of infection (Haase *et al.*, 1982, Harris *et al.*, 1984) is likely to force any processes occurring at low frequency *in vivo* to much higher rates, with the possibility of allowing interactions which would be blocked *in vivo* due to low affinity.

3.2.4 Selection of LTRs for functional analysis

The sequencing of the EV-1 virus showed a large number of LTR variants. However, simply by analysing the sequence it is impossible to ascertain whether these are functional promoters. None of the sequences generated were obviously defective in that they all contain conserved TATA boxes and sequences with homology to transcription factor binding sites. Functional analysis was carried out to answer two separate questions; firstly are the variants active as promoters and secondly are there significant differences in their activity.

For these studies a representative group of the *ex vitro* LTRs were selected for analysis; LTRs 1,18,19,28,30,44, and 48. These LTRs together with the 1514 LTR are illustrated in Fig. 3.4. LTRs 18,44 and 48 bear closest similarity to the *ex vivo* sequences. The repeat boundaries of LTR 18 match those found *in vivo*, while the repeats in LTR 44 are slightly extended when compared to those seen *in vivo*. LTR 48 is homologous to the repeatless LTR forms found *in vivo.*

While only a single repeat type is observed in the *ex vivo* LTR population the *ex-vitro* LTRs show a greater number of repeat patterns. These distinct structures are represented by LTRs 28,19 and 30: LTRs 19 and 30 show similar but not identical repeat structure, with LTR 1 possessing no repeats.

The 1514 LTR was included in this study as it is the only MVV LTR to be characterised functionally. Comparison of the EV-1 sequences with 1514 can be used to address the significance of the divergence between these two viruses.

To summarise, the LTRs used in this study can be divided into two groups on the basis of their similarity to *ex vivo* sequence. LTRs 18,44 and 48 can be classified as being *ex vivo* like in that they most closely match the sequence and repeat boundaries of the LTRs present in infected animals. LTRs 1,19,28 and 30 differ in that they all lack the insert at position 88 and in the case of LTR 28 possess a repeat not represented in *ex vivo* sequences.

3.3 Determination of relative LTR activities

3.3.1 Experimental design

In attempting to assess the significance of any variation in LTR activity seen in these experiments the possible impact of systematic errors during transfection analysis must be addressed. These can be divided into two different categories, variation in transfection efficiency and quantification of reporter gene products within transfected cells.

Transfection efficiency can vary according to several parameters; variation in the amount of DNA transfected (inaccurate DNA quantitation), contaminants in the DNA which inhibit transfection or are toxic to cells and cell passage number. The use of a second plasmid with an assayable product should allow correction for most of these errors. In these experiments the ß-gal gene under the control of the SV-40 promoter and enhancer was used.

FIGURE 3.4 Alignment of the U3 regions of the LTRs used in this study

The eight LTRs which were selected for analysis are aligned above. Potential transcription factor binding sites are shown above the sequences. The question marks denote sites which do not show full homology to the reported consensus sequence (Faisst & Meyer 1992). The TATA box proximal AP-1 site is consensus only in the 1514 sequence. Sequences which were shown to footprint in the 1514 LTR (Gabuzda *et al.*, 1989) are underlined.

Reduced transfection efficiencies due to variation in quantity or quality of the DNA under investigation will also lead to variation in transfection efficiency of the second plasmid. In these studies DNA purity and concentration was checked by spectrophotometer and DNA integrity confirmed by gel electrophoresis. Throughout these studies the level of ß-gal activity was consistent within different transfection experiments suggesting no major variation in DNA quality or transfection efficiency.

The second set of errors may occur during preparation of cell extracts and quantification of reporter gene products. Any technique for the production of cell extracts for enzyme assay is likely to give some variation in efficiency. As a second control enzyme, ß-Gal, was present in addition to CAT it is assumed that both of these proteins should be extracted from cells with a similar efficiency during the freeze thaw process. Since a second enzyme was used the final measure of promoter (LTR) activity is given by the CAT:ß-gal ratio for each cell extract in an experiment. A second point in relation to the quantitation of reporter gene activity is that for these studies enzyme activities were assayed directly rather than using the products of the transfection vectors as substrates for ELISA assays. Any enzyme assay must be performed in the linear range to ensure valid results. For the ß-Gal assays this does not represent a problem as there tended to be no significant variation in ß-gal activity between extracts within each transfection. In the case of the CAT assay, however, different constructs did show a several fold variation in activity. For this reason the phase extraction assay using n-butyryl CoA was used, this assay being linear for CAT concentrations varying by three orders of magnitude (Seed & Sheen 1988). None of the constructs used in this study showed variation of this magnitude allowing all samples to be assayed in parallel without dilution.

One potential drawback of using a second assay plasmid to control for transfection efficiency is that the two plasmids may interfere with each other (Farr & Roman 1991). This effect may at least in part be accounted for by competition between the plasmid promoters for limiting cellular transcription factors. In order to test that the pCAT LTR and ß-Gal vectors were compatible, chondrocytes were transfected with 5µg of pCAT 1514 (which contains the 1514 LTR sequence) and 5µg of either pSV2ß-Gal or pUC18. For these experiments the CAT activity in the presence of the ß-Gal plasmid was compared to that seen in its absence. When expressed as a percentage, the activity in the presence of pSV2ßGal was 98 +/- 11% of the that when no SV-40 promoter enhancer sequences were present. This represents data from four independent transfections.

3.3.2 LTR Activity in transfection assays

The relative activities of these eight LTRs were analysed by transfection into chondrocytes growing in 5% FCS. Transfections were performed 44 hours prior to preparation of cell extracts by freeze thaw lysis. Both CAT and ß-Gal activities were determined for all extracts using the ß-Gal activity to correct for transfection efficiency. For this analysis the activity of the EV-1 LTRs was expressed as a percentage of 1514 activity.

The data from these experiments is shown in Fig.3.5. All of the EV-1 derived LTRs are active and show activities equal to or below that of the 1514 LTR. These data show significant variation in activity within the group of LTRs tested. Under these assay conditions the activity of the pCAT12 vector was >1% of the 1514 LTR.

The main question to be addressed is the effect of sequence and structural variation on LTR activity. Those LTRs with similar repeat structures but slightly varying boundaries, i.e. LTRs19 & 30 and 18 & 44, show the same relative activity ($p>0.05$) suggesting that the exact positioning of repeat boundaries has little effect on LTR activity.

The effect of the repeat in modulating the activity of the LTR in comparison with the repeatless version is seen to be dependent on the entire LTR sequence and not just the repeat region itself. This is shown by comparing the activity of LTR1 with LTRs 19 & 30 and LTR48 with LTRs 19 & 44. LTRs 18, 19, 30 and 44 have closely related repeat regions but the introduction of these repeats in one case reduces LTR activity (LTRs 19 and 30 versus LTR 1, $p<.0001$) and in the other elevates activity (LTRs 18 and 44 versus LTR 48, $p<.0005$). The activities of LTRs 19 & 30 are equivalent to LTR 48 (p values of 0.23 and 0.36 respectively).

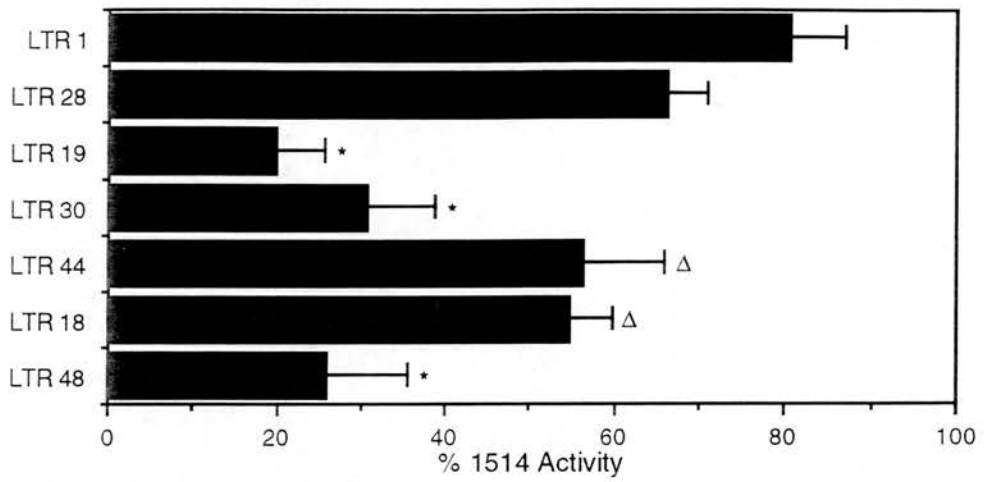
The most obvious distinction between LTRs 48 and 1 is the presence of the 12 base pair insert in LTR 48 (and the related LTRs 18 and 44). However, these two LTRs also differ at a number of other positions due to point mutations. The relative activities of LTR 1 and 48 suggest that while the context of LTR 48 allows a positive action for repeat structures its activity is lower than that of LTR 1. The single LTR with an upstream repeat analysed in this study, LTR 28, shows an activity intermediate between LTR 1 and LTR 48.

Taken together, this data would support the proposition that promoter activity is not determined by having a given number of transcription factor binding sites within a promoter, but rather is dependent on the relative positions of these sites. This is seen in comparison of potential transcription factor sites within different sets of repeats. While LTR 44 repeats

The raw data used to generate figure 3.5 is shown below. Data is given as percent 1514 activity.

	Expt.	1	2	3	4	5	6	mean	SD
LTR	1	89	76	86	81	71	80	80.5	6.5
	18	51	57	54	62	55	49	54.7	4.6
	19	28	13	20	19	24	15	19.8	5.6
	28	71	67	70	66	60	63	66.2	4.2
	30	23	19	34	40	34	35	30.8	8.0
	44	54	45	72	60	58	48	56.2	9.6
	48	15	36	24	37	15	28	25.8	9.7

Figure 3.5 Functional Comparison of EV-1 LTRs



Activity of EV-1 LTR sequences in pCAT12 vector expressed as %1514 activity. All points are the average of six independent transfections with two plasmid preparations. Data is expressed as mean plus one standard deviation.

* - Denotes sequences significantly different from LTR 1 ($p<0.02$)

Δ - Denotes sequences significantly different from LTR 48 and LTR 1 ($p<0.05$)

Data was analysed using a two sample t-test.

essentially the same region as LTRs 19 and 30, it shows significantly higher promoter activity and an elevation in activity from its repeatless counterpart LTR 48. In contrast, LTRs 19 and 30 show a much reduced activity when compared to LTR 1 even though these repeats cover regions believed to bind transactivating factors such as AP-1. Indeed LTR 30 differs from LTR 1 at only a single base, an A to G conversion within the repeated sequence. This would suggest that it is the presence of the repeat and not point mutations which are the cause of the decreased activity of LTR 30 in comparison with LTR 1, which was highly consistent between different DNA preparations.

In relation to the *ex-vivo* like LTRs 18/44 and 48, the elevation in activity would suggest the presence of binding sites for transactivating factors within the repeated region. An alternative explanation is also possible in that the repeat removes a negative regulator, present in the upstream region, sufficiently distant from the initiation complex around the TATA box to dilute its effect. A full explanation of these observations on LTR activity will be dependent on a description of the transcription factors which may be interacting with the LTR within the cell.

3.4 Identification of potential transcription factor binding sites in the EV-1 LTRs

The transfection assays described above illustrate the effects of repeat structure on LTR activity. From these experiments, it appears likely that there may be a binding site for a transcriptional activator within the region outlined in Fig 3.6. This Figure also shows this region from two other MVV viruses; SA-OMVV and KV1772.

This region contains the only site in the MVV LTR which interacts with a characterised transcription factor. In footprinting and gel retardation experiments using the 1514 LTR (Gabuzda *et al.*, 1989, Shih *et al.*, 1992) the TATA box proximal consensus AP-1 site has been shown to interact with cellular Fos and Jun proteins (Shih *et al.*, 1992). This is in contrast to the other degenerate AP-1 sites present within the 1514 LTR. This consensus AP-1 site is mutated in the EV-1 LTRs (Fig. 3.1/2/6). The loss of the TATA box proximal consensus site is observed in all EV-1 LTR types, with none of the *ex vivo* or *ex vitro* LTR sequences showing a consensus site. This argues against the loss of this consensus sequence being a PCR artifact.

FIGURE 3.6 Alignment of AP-1/AP-4/AP-1 regions of MVV isolates

AP-1 consensus TGA$^{G}/_{C}T^{C}/_{A}A$

AP-4 consensus $^{C}/_{T}$CAGCTG$^{C}/_{T}$GG

```
100                                 130                 150
     AP-1          AP-4          AP-1
GCGCTAAGTCATGTAGCAGCTGATGCTTGAGTCATAACCGCAGATGTAAA  KV1514
....G.....................G..A.....................  KV1772
....AT.........T............A.............AT......  EV-1 ex-vivo
....A.........T............A.........A..AT......  EV-1 ex-vitro
A......A.......T............AG............AT......  SA-OMVV
```

Sequence of region which shows modulation of LTR activity. Sequence is compared between MVV isolates KV1514 (Sonigo *et al.*, 1985), KV1772 (Andresson *et al.*, 1993), SA-OMVV (Querat *et al.*, 1990) and EV-1 (Sargan *et al.*, 1991).

The region contains putative AP-1 and AP-4 recognition sequences. Comparison of these sequences with the consensus sequences (Faisst & Meyer 1992) shows that while both KV1514 and KV1772 sequences have full AP-1 consensus sequences (underlined regions) the sites within the other sequences are degenerate.

Nucleotide positions are the same as Fig.3.1 and 3.4.

The loss of the consensus AP-1 site in the EV-1 LTRs is in contrast to the 1772 (Icelandic) neurovirulant isolate derived from the 1514 virus. In this virus the TATA box proximal site (position 129) is mutated from consensus. However, the site at position 106 (Fig 3.6) is back mutated to a consensus AP-1 site. This observation would appear to suggest that the presence of a consensus AP-1 site, within this region, is important for LTR function, at least in the context of the 1514 LTR and those derived from it. The SA-OMVV virus (a South African isolate) also shows the loss of the consensus AP-1 site. This virus is a distinct geographical isolate to EV-1 suggesting that LTR regulation may have diverged between distinct geographical isolates.

Comparison of the AP-1 sites present in EV-1 and 1514 (Fig. 3.7) shows that while the EV-1 LTRs have a number of degenerate AP-1 sites none of these match the consensus sequence. AP-1 binding is not absolutely dependent on the presence of a full consensus sequence (van Lint *et al.*, 1991, Jain *et al.*, 1992). However, variation from the optimum (consensus) site does result in a reduction of site affinity. None of the 1514 AP-1 sites with the exception of the consensus site at position 129 has been shown to bind AP-1 (Gabuzda *et al.*, 1989). Only in one case do mutations in the EV-1 sequences generate a sequence with closer homolgy to the AP-1 consensus; position 56 in LTRs 1, 28, 19 & 30. This sequence matches that known to bind AP-1 with low affinity in the IL-2 promoter (Jain *et al.*, 1992). It should, however, be noted that the site found at this position in *all* the *ex vivo* LTRs is that present in LTRs 18, 44 & 48 which is even further from consensus than the sequence which failed to footprint in 1514. The observation that repeats of the region outlined in Fig.3.6 can result in elevation of transcription suggests that there is a functional site for a cellular transcription factor within this region. The presence of the near consensus AP-1 site (position 129) in this region suggests that this site may be responsible for this observation.

While it is impossible to discount AP-1 interactions with the EV-1 LTRs simply by sequence comparisons two preliminary conclusions may still be drawn. Firstly that the EV-1 LTRs do not contain a high affinity AP-1 site. If the factor is binding it is doing so to a low affinity non-consensus site. The second, and related point is that, these two viruses EV-1 and 1514 may not be regulating transcription from the LTR in an identical manner. The final result may be similar but the underlying mechanism different.

Figure 3.7 Potential AP-1 Sites within the EV-1 and 1514 LTRs

Position 56	1514	ATAGTCA	
	EV-1	AGAGTCA	LTRs 1,28,19 & 30
		AAAGTCA	LTRs 18, 44 & 48
Position 66	1514	TGACACA	
	EV-1	TGACACA	LTRs 18, 30, 44 & 48
		TGACATA	LTRs 1, 19 &28
		TGACATC	LTR 19b (second repeat)
Position 106	1514	TAAGTCA	
	EV-1	AAAGTCA	LTRs 1, 19, 28 & 30
		ATAGTCA	LTRs 18, 44 & 48
Position 129	1514	TGAGTCA	
	EV-1	TAAGTCA	All EV-1 LTRs
AP-1 consensus		TGAGTCA C A	

AP-1 sites are shown from EV-1 and 1514 LTRs. Positions of sites are as shown on Fig.3.4. In 1514 the only site which binds AP-1 is at position 129 (the TATA box proximal site), all the other degenerate 1514 sequences do not bind the factor (Gabuzda *et al.*, 1989). The EV-1 sequences are compared to the 1514 sites. Base changes in the EV-1 sites are underlined. In no case do the differences from the degenerate 1514 AP-1 sequences generate consensus recognition sites in the corresponding EV-1 sequences. These sequences are discussed in relation to the *ex vivo* LTRs in the main text.

In addition to the DNase footprint over the AP-1 site, two other sites have been shown to generate specific footprints in the 1514 LTR (Gabuzda *et al.*, 1989): one at position 75 (in the second repeat) the other at position 90 (Fig.3.4). The 1514 sequences footprinted are compared with those present within EV-1 in Fig.3.8. The footprint in 1514 at position 75 is present as a single copy at the end of the 53 bp repeat and extends beyond the repeated sequence. The EV-1 LTRs also have a single copy of this region. This region is partially conserved between the two viruses (Fig.3.8) and so may be functional in both EV-1 and 1514 LTRs.

The other 1514 DNase footprint (position 90 Fig.3.4) is over a region which is not present in all the EV-1 LTRs. This site at position 90 in the 1514 LTR only has a potential homolog in the *ex vivo* LTRs (Fig.3.2) and *ex vitro* LTRs18, 44 and 48, the *ex vitro* LTRs with closest similarity to those found *in vivo*. Although this site is not identical in 1514 and these *ex-vivo* EV-1 LTR types it is homologous (Fig.3.8). This sequence is deleted in the EV-1 LTRs 1, 19, 28 and 30. In relation to the transfection data this sequence appears to have a negative effect on the transcription from the LTR, LTR 1 compared to LTR 48 (Fig.3.5).

The MVV LTR contains a number of other sequences which are potential targets for transcription factor binding. Most of these are degenerate with low homology to consensus sequences. It is problematic to assign functionality to these sequences for two reasons. Firstly consensus sequences are often just 'best fits' and do not represent the only sequences capable of binding a factor (Mermod *et al.*, 1988); they are useful only as a rough guide to indentify putative high affinity binding sites. Secondly flanking regions can play a critical role in determining a site's affinity for a given transcription factor (Ryseck & Bravo 1991, Nielsen *et al.*, 1994).

One example of these degenerate transcription factor sites within the LTR is the non-consensus AP-4 site at position 116 (Fig.3.3). While there is no evidence that this site binds AP-4 it has been proposed as the target for a factor which recognises both this sequence and the down stream proximal AP-1 site (Neuveut *et al.*, 1993). This factor appears dependent on the presence of both the AP-4 and AP-1 sequences.

Two further sites of interest within both the 1514 and EV-1 LTRs are the consensus E-box sites (CANNTG) (Faisst & Meyer 1992) at positions 73 and 142 in EV-1 and positions 73 (within both copies of the 53 bp repeat) and 142 in 1514. This site is in close proximity to the 1514 footprint site (position 76) although not underlying it (Fig.3.4). The E-box transcription

Figure 3.8 Potential Transcription Factor Binding Sites in the EV-1 LTR

Position 75

1514 Footprint	AAATGTAACCGCAAGTTCT	
	----------GCGC	LTR 1 & 28
	--------T------GCGC	LTR 19 (1st repeat)
	----------GCGCA	LTR 19 (2nd repeat)
	----------GCGC	LTR 30 (1st)
	--G---------GCGC	LTR 30 (2nd)
	----------TTCTG	LTR 18 (both repeats)
	----------	LTR 44 (1st)
	----------GCGC	LTR 44 (2nd)
	----------TTCTG	LTR 48

Position 96

1514 Footprint	TTTTTTG	
	--AAA--	LTRs 18, 44, 48

Comparison of footprinted regions of the 1514 LTR with the analagous regions of the EV-1 LTRs. The *ex vivo* LTR sequences match LTRs 48 & 18 for the sequence at the position75 footprint. Only LTRs 18, 44 & 48 of the *ex vitro* LTRs contain the insert which contains the 1514 position 96 footprint. All *ex vivo* LTRs contain this sequence.

factors are members of the basic Helix-Loop-Helix (bHLH) family (reviewed by Garrel & Campuzano 1991). Like the products of the *fos* and *jun* gene families these proteins form dimers so modulating their transactivating and sequence recognition abilities. E-box factor binding sites have been described as both positive and negative regulatory elements depending on cell type and differentiation stage. The range of E-box binding factors expressed by any given cell type will depend on cell lineage and activation state. These two E-boxes are in addition to the E-box site within the degenarate AP-4 site which makes up part of the AP-4 recognition sequence.

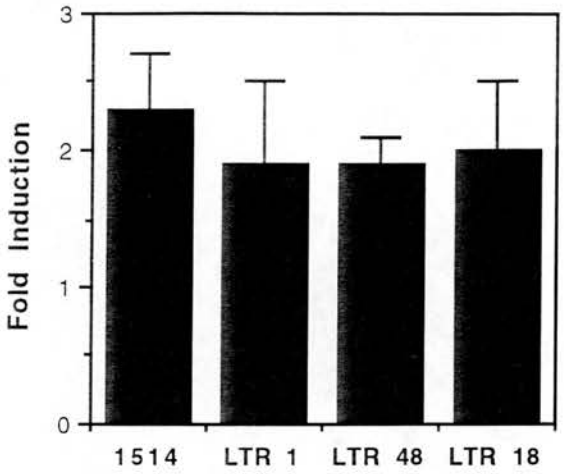
3.5 Serum induction of MVV LTRs

In addition to the comparisons of steady state LTR activity, it was decided to analyse the induction potential of the different LTRs. A number of transcription factors, including AP-1, are known to be upregulated upon serum stimulation of cells. Differences in the induction potential of the LTRs may therefore reflect variations in transcription factor binding.

For these experiments the cells were transfected and then placed in 0.5% FCS. This was replaced with either fresh 0.5% FCS media or media supplemented with 10% FCS 12 hrs prior to harvesting. It was not possible to withdraw completely serum from the chondrocytes after transfection as this leads to extensive cell death. For this experiment only LTRs 1, 18, 48 and 1514 were assayed.

The results from these experiments are shown in Fig.3.9. Under these conditions all four LTRs show identical levels of induction (approximately 2 fold). This similarity in response may be due to the same transactivating factors acting on each LTR type or alternatively different factors with similar induction/transactivating potential. This is particularly relevant in comparison between the 1514 and EV-1 LTRs in relation to the apparent lack of a consensus AP-1 site within the EV-1 LTRs. A further point is the role of inducible versus constitutive transcription factors in determining LTR activity within this system. The exact nature of the chondrocytes' response to serum induction is unknown, although these cells do show an inducible AP-1 activity (Sections 4.6.1 and 4.7). While it is possible to assign the induction observed to inducible DNA binding proteins an alternative explanation is that the observed induction is due to an upregulation of other components of the transcription complex, such as the basal factors which assemble on the TATA box binding protein. Thus the effect observed

FIGURE 3.9 Serum induction of LTRs



The serum induction of the four LTRs 1, 18, 48 and 1514 was compared. Induced cells were grown in 10% FCS. Induction is measured over the activity of LTRs in cells growing in 0.5% FCS. The data for all LTRs is the mean of four independent transfections, with the exception of the 1514 data which is from 5 transfections.

is not promoter specific but could represent a general activation of transcription. This supported by the increase in expression from the ß-Gal plasmid, both vectors showed an equivalent rise in activity.

3.6 Discussion

3.6.1 Is the LTR of MVV virus under selective pressure *In Vivo*?

The EV-1 LTRs isolated from infected animals show reduced diversity when compared to those seen *in vitro*. One possible explanation, that many of the *ex vitro* LTR types are non-functional, is countered by the observation that all the *ex vitro* LTRs tested by transfection analysis were functional and could drive transcription. These two observations, the reduced diversity of LTR types seen *in vivo* and the lack of non-functional LTR types in the *ex vitro* population, suggests that the MVV LTR may be under selective pressure *in vivo*. Comparing the LTR populations in the two animals from which the *ex vivo* LTRs were derived it is clear that in both animals the same LTR type has been selected. This LTR appears to be most closely related to the *ex vitro* LTR 48.

Examining the *ex vivo* LTR sequences presented in Fig.3.2 then sheep 848, the chronically infected animal, shows only repeatless sequence in both blood and lymph. These LTRs differ at a larger number of sites due to point mutation so a number of these alterations may represent true population variation. As the only variation between these sequences is accounted for by point mutation, and not introduction of repeats, it is possible that a single progenitor LTR type was present at the initiation of infection.

In sheep 649, there was only a single LTR type early in infection (day 9). In the day 9 population only 1 out of 10 LTRs (LTR 13) shows variation. This mutation, a TC inversion at position 225 (Fig.3.2), might be assigned to PCR error suggesting that only a single LTR sequence was selected from the infecting population (Fig.3.1). At later time points, 15 days and 5 months post infection, there is evidence for the generation of repeat structures within this LTR, although only a single repeat type appears to be tolerated. As well as repeats, the later time points from sheep 649 show an accumulation of point mutations which may illustrate sequence divergence within the viral population. An important point to note is that these mutated forms accumulate later in infection and are not present at the early time point. This

lack of sequence diversity early in infection suggests a possible bottleneck in LTR evolution within the animal. The error rate for the PCR used to generate these sequences was low, estimated at less than 1 in 2400 bases, for sequences derived from LTR 7. Thus most of the sequence variation observed can probably be described as genuine rather than as an artifact of the PCR amplification.

Two mechanisms may account for this apparent *in vivo* selection. Firstly, of all the LTRs seen *in vitro* only a single form may be functional *in vivo,* i.e. an LTR 48 derivative. A second explanation is that early in infection only a small population of infected cells escape the immune response. In this scenario, the selection of the founding LTR population is random and independent of LTR sequence. As only two infected animals were analysed this second argument cannot be discarded, although given that there is a large population of distinct *in vitro* variants it is surprising that in both infected animals the same progenitor LTR sequence appears to have been selected.

The data from the transfection experiments suggests that the variations observed in LTR structure affect the rate of transcription from the LTR. Combining the results of the transfection analysis with the variation in LTR type seen *in vivo* then it appears that early in infection (day 9, sheep 649) only LTRs with low rates of replication are selected. The reason why LTR 48 like structures are selected over LTRs 19 and 30 is unclear but suggests that sequence considerations also play a part and that not any LTR with a low promoter activity will be selected. This suggests that the insert at position 90 while not a requirement for *in vitro* growth is required *in vivo*. At later time points repeat variants appear to be tolerated (Sheep 649 day 9 and 5 months) although only a single repeat type is represented. This would suggest that these repeat-bearing LTRs are derived from the day 9 LTR rather than representing virus not detected at the earlier time points.

The LTR population found in sheep 848 (the chronicaly infected animal) suggests that LTR variants containing repeats are not a requirement for on going infection and virus replication late in disease can be maintained by a similar LTR type to that found at the onset of infection. The issue of LTR selection during the PCR process should be mentioned. While there is unlikely to be linkage between the sequences used to generate primers for the amplification of LTR sequence and the U3 promoter regions within the LTR, it is still true that the LTR population described is not an exact match to that actually present within the animal.

3.6.2 Possible significance of variations in LTR activity

While the effects on LTR activity caused by sequence variation were not large they may have implications for the regulation of viral replication. This statement is dependent on the observed differences in LTR activity holding true for the transcription of the entire viral genome during *in vivo* infection. Extrapolation of these results to the situation *in vivo* is hampered by several observations. These involve the validity of comparing simplified *in vitro* systems with the process of viral replication *in vivo*.

The first issue is the cell type that is used to analyse LTR activity in these studies is not the cell type predominantly infected *in vivo* (Brahic *et al.*, 1981, Narayan *et al.*, 1982, Narayan *et al.*, 1983 & Gendelman *et al.*, 1985). While MVV virus will productively infect chondrocytes *in vitro* it is unknown whether or not this cell type is productively infected *in vivo*. The target cells for MVV infection *in vivo* are primarily those of the monocyte and macrophage lineage and their precursors in the bone marrow. Productive infection *in vivo* is not limited to these cell types. There is some evidence for other, epithelial and fibroblast like cells also being productively infected, at low frequency, *in vivo* (Chapter 1.4.4., C. Zink personal communication). Previous studies on the 1514 LTR have made use of a variety of different cell types. These studies have shown that while the absolute activity does vary, different cell lines reflect the same general trends. This would support the proposition that the differences in LTR activity seen within these studies would be reflected over a variety of cell lines and possibly also *in vivo*.

While the cell type under study may not affect the general conclusions on variant activity it should still be pointed out that these two cell types, chondrocytes and macrophages, are likely to express overlapping but distinct sets of transcription factors. Some, such as AP-1, are likely to be present in both. However the exact composition, i.e. which members of the *jun* and *fos* families the cells express, is likely to vary. This variation in composition may have an effect on the transactivating properties of the complexes generated.There is also the possibility of lineage restricted factors which are present in the monocyte/macrophage lineage but not chondrocytes. Thus while the variation in activity of the different LTR structures used in this study may be genuine the magnitude of functional variation may differ *in-vivo*.

A second issue is the absence of viral transactivators in the comparison of LTR activity

described here. In normal MVV infection the transactivator protein Tat will also be regulating the rate of viral transcription. The MVV Tat protein appears to function in a distinct manner to that of the other lentiviral Tat proteins, with the possible exception of the CAEV and FIV proteins, due to the lack of a TAR region in MVV (Chapter 1.5.3). The MVV Tat protein therefore appears to be acting solely through cellular transcription factors which bind to the LTR (Gdovin & Clements 1992, Neuveut *et al.*, 1993). This effect could be mediated by one of two mechanisms. The MVV Tat protein may be directly transactivating cellular genes so raising the levels of transcription factors which can activate transcription from the LTR. This effect has been described in relation to the cellular *c-jun* gene (Neuveut *et al.*, 1993). A second point of action is the transcription complex assembled on the LTR. The MVV Tat protein could be interacting with cellular transcription factors so directing its transcription activation domain (Carruth *et al.*, 1994) to promoters. This mode of action is seen in the HTLV transactivator Tax (Zao *et al.*, 1991, Zao *et al.*, 1992, Adya *et al.*, 1994) and the HBV X protein (Maguire *et al.*, 1991). It should be noted that these two modes of action are not mutually exclusive. If the MVV Tat protein is acting in this way, via interactions with cellular transcription factors, it is likely to trigger a general cellular activation or maintenance of an activated phenotype following normal stimulation of the infected cell *in vivo*. If so, the Tat protein is relying on interactions with cellular factors for targeting to the LTR it will also be targeted to any other cellular promoter or enhancer binding that factor or factors. MVV Tat may therefore boost viral replication not by greatly elevating the rate of transcription from the viral LTR but rather by keeping the infected cell in an activated state. This would maintain high levels of cellular factors so allowing viral transcription to continue for a prolonged period. This may explain why the MVV LTR shows only weak transactivation in response to Tat; the LTR does not represent the primary, direct, target for transactivation. The absence of Tat within these studies is most likely to show itself in a general reduction of activity rather than as a gross differential effect between LTRs, with the possible exception of the EV-1 LTRs versus 1514.

Differences between these two sets of LTRs, in relation to their Tat responsiveness, may arise in two ways. Firstly the sequence divergence between the two strains may have altered the transactivation potential of the two Tat proteins. Secondly MVV Tat has been proposed to be to acting primarily through the consensus AP-1 site and the adjacent AP-4 like sequence within the 1514 LTR (Gdovin *et al.*, 1992 & Neuveut *et al.*, 1993). The lack of a consensus AP-1 site in the EV-1 LTRs may therefore result in an alteration of Tat

transactivation potential or altered target for Tat activity within the EV-1 LTRs.

Given these problems around variations in transcription factor expression between chondrocytes and macrophages and the possible effects of Tat on the LTR sequences studied here it is still possible to draw some tentative conclusions. The first is that a certain LTR structure, and possibly activity, is prefered *in vivo*. This would appear to be a structure which gives a relatively low level of transcription; the most active LTR types do not appear to be selected. In relation to the viral lifecycle *in vivo* this observation fits in with the idea of immune evasion by restricted replication or latency. There is insufficient antigen production within the infected cell to trigger immune effector cells (most importantly CD8 cytotoxic cells). This would lead to the survival and dissemination of infected monocytes. Activation of cells could then lead to a burst of viral replication before killing by immune effector cells. This 'Trojan Horse' model for the spread of virus is dependent on the infected cell being able to escape immune surveilance up until the point the cell is activated so allowing maximal production of progeny virus before killing. This slow spread of virus in MVV infected sheep is supported by the observation that there is little or no detectable virus within the blood of infected animals. This type of slow spread may be of more importance to MVV than the immunodeficiency viruses, where suppression of the immune sytem eventually allows almost unchecked viral replication at the end stage of disease. In MVV there appears to be no major defect in the cell mediated immune response right up until death. However, there is a reduction in IgG2 levels in MVV infected animals which may supress antibody dependent cell cytotoxicity and indicate a more subtle disruption of the immune system than is seen in the immunodeficiency viruses. Tight regulation of transcription is thus central to the survival of MVV within the animal. It should be stated that this regulation of transcription and expression of viral protein is under the control of sequences other than the LTR. The Tat and Rev proteins will also play a role. However, the U3 promoter sequences play a pivitol role in that they probably represent the on/off switch for viral transcription and control the accumulation of these early proteins which may trigger the switch to the production of infectious virus when they accumulate to a sufficient level.

The second point raised by the sequence comparisons and functional analysis of LTRs in this chapter is whether the EV-1 and 1514 have truly identical transcriptional control strategies. The comparison of activities shows that the *ex vivo* EV-1 LTRs show a lower activity than the 1514 LTR. While this difference was not huge it was reproducible. The absence of a

consensus AP-1 site in the AP-1/AP-4/AP-1 region of the EV-1 viruses when compared to its maintenance in the 1772 and 1514 viruses suggests that while the 1514 and EV-1 LTRs have very similar LTR structures there may be some differences in the fine tuning of the mechanisms controlling transcription.

As the consensus AP-1 site in the 1514 LTR has been proposed as the principle regulator of LTR induction, and given the importance of induction in the lentiviral lifecycle, it seems unlikely that the EV-1 LTRs have completely lost the ability to bind this factor. As LTR 48 can be taken as the best match to the 1514 LTR of all the EV-1 sequences and this sequence responds in a positive manner to duplication of the AP-1/AP-4/AP-1 element (LTR 18) it seems likely that this region of the EV-1 LTRs does contain a *cis* acting positive regulatory element. The best candidate for this sequence is the degenarate AP-1 site at position 129 (Fig.3.7). The role of this sequence will be more closely examined in the next Chapter.

CHAPTER FOUR

4.1 Introduction

In the previous chapter experiments with the EV-1 LTR variants suggested that the AP-1/AP-4/AP-1 region (Fig.3.6) in the EV-1 LTRs was a possible target for a transcriptional activator. Duplication of this region is, in the correct context, capable of elevating the activity of the LTR (Fig.3.5). The sequence most likely to be mediating this effect is the TATA box proximal AP-1 site. However, this AP-1 site does not match the consensus AP-1 site found at the same position in the 1514 LTR. In the 1514 LTR this AP-1 site has been described as one of the principal regulatory sequences (Gdovin *et al.*, 1992 & Shih *et al.*, 1992). The EV-1 and 1514 sequences at this site diverge by a single base transition (G to A). As this site differs at only one position from the functional sequence seen in the 1514 LTR and the mutation is not one that is known to abolish binding, it was decided to compare directly the binding characteristics of these two sequences. A list of the oligonucleotides used in this Chapter is given in Fig.4.1a and their position of the MVV derived sequences within the LTR in Fig.4.1b.

The variation between the EV-1 and 1514 sequence in this region may alter its function in one of two ways. While AP-1 may still bind to the EV-1 sequence it may do so with an altered, probably reduced, affinity. The altered AP-1 site in EV-1 may thus preferentially bind a distinct set of AP-1 dimers. This may result in subtle alterations in the response of the two viruses to monocyte activation and differentiation *in vivo*. The second possibility is that AP-1 binding to this site has been abolished by the mutation in the recognition sequence. If this is correct, and AP-1 is no longer binding to the EV-1 LTRs, then it raises further questions. Primarily, given that the consensus AP-1 site appears to be required for inducibility of the 1514 LTR (Hess *et al.*, 1989); how are the EV-1 LTRs compensating for the apparent loss of AP-1 binding to this site?

In order to test the binding properties of the degenerate TATA box proximal AP-1 site present in the EV-1 LTRs, double stranded oligonucleotides corresponding to the EV-1 and 1514 regions were compared in gel retardation assays. This assay takes advantage of the stability of protein/DNA complexes under non-denaturing conditions in native polyacrylamide gels (reviewed in Lane *et al.*, 1992).

The protocol used for these studies was fully described in Chapter 2. Unless stated otherwise, all experiments were performed using 5μg of nuclear protein extract prepared from chondrocytes growing in 5% FCS. This extract was initially incubated with poly dI-dC.dI-dC to

FIGURE 4.1a Synthetic oligonucleotides

Control AP-1
TTCCGGC**TGACTCA**TCAAGCG
AAGGCCG**ACTGAGT**AGTTCGC

1514 AP-1
TGCT**TGAGTCA**TAACC
ACGA**ACTCAGT**ATTGG

EV-1 AP-1
TGCT**TAAGTCA**TAACC
ACGA**ATTCAGT**ATTGG

1514 oligo
TGATGCT**TGAGTCA**TAACCGCA
ACTACGA**ACTCAGT**ATTGGCGT

EV-1*ex vivo*
TGATGCT**TAAGTCA**TAACCGCA
ACTACGA**ATTCAGT**ATTGGCGT

EV-1*ex vitro*
TGATGCT**TAAGTCA**TAACCACA
ACTACGA**ATTCAGT**ATTGGTGT

Bases underlined are those which differ between the EV-1ex/EV-1 and 1514 sequences.

VSS oligo
GAGTCATAACCGCA
CTCAGTATTGGCGA

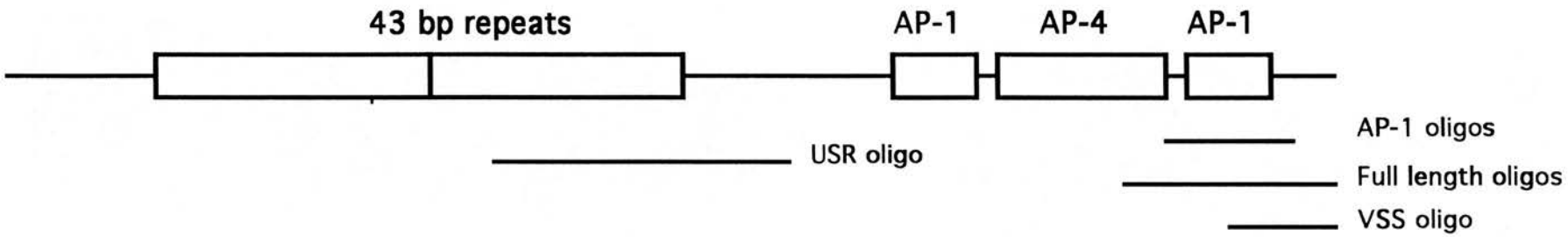
Upstream Region (USR)
TCAGGA**TGACACA**GCAAATGTAACCGCAAGTTCTGCTT
AGTCCT**ACTGTGT**CGTTTACATTGGCGTTCAAGACGAA

Tα4
CCCCCCAACCGCAGGTGCAG
GGGGGGTTGGCGTCCACGTC

X-link VSS
CAGTCATAACCGCA
GTCAGTATXGGCGX X=5 Bromo dU

Sequences shown in bold are consensus or putative AP-1 sites

Figure 4.1b



The relative positions of the varios oligonucleotides are outlined against the 1514 LTR.

remove nonspecific DNA binding proteins. After this incubation competitor oligonucleotides were added at 1000x fold excess over radiolabelled probe. Finally the radiolabelled probe was added and the reaction incubated for a final 10 minutes prior to loading onto the gel. The complexes formed in these reactions were resolved on 4% non-denaturing polyacrylamide gels. Retarded probe was detected by autoradiography of dried gels.

4.2 Comparison of EV-1 and 1514 AP-1 sites

The TATA box proximal AP-1 site in MVV 1514 has been shown to bind the transcription factor AP-1 (Shih *et al.*, 1992, Gabuzda *et al.*, 1989). The AP-1 sequence in the 1514 LTR fully matches the consensus site but the site at this position in the EV-1 LTRs diverges from the consensus sequence. In order to examine whether this site was functional in the EV-1 LTRs oligonucleotides covering this area were compared. The two sequences used (1514 AP-1 and EV-1 AP-1) and the results of gel shift experiments using these double stranded oligonucleotides as probes are shown in Fig.4.2a/b. It is important to state that the sequence used here in the EV-1 AP-1 oligonucleotide is found in all the EV-1 LTR sequences, both *in vitro* and *in vivo*. The chondrocytes used to generate the cell extracts used in these experiments are known to express a Jun-containing AP-1 sequence binding activity (Section 4.6.1).

A control AP-1 sequence was used to confirm that any shifts observed with the MVV probes were due to AP-1 binding and not an additional transcription factor site in this region. The control probe from the human collagenase promoter gives a single retarded band (lane 9 Fig.4.2a). The protein in this complex is competed by an excess of both cold control AP-1 (lane 11) and 1514 AP-1 (lane 10) oligonucleotides. No competition is seen using the EV-1 AP-1 sequence (lane 12). When using the 1514 AP-1 sequence as probe a retarded band is observed (lane 5). The retarded complex is competed by both 1514 AP-1 and control AP-1 oligonucleotides (lanes 6 and 7 respectively). It is clear that the complex binding to the1514 AP-1 sequence is doing so with a lower affinity than that binding to the control AP-1 sequence (lane 5 versus lane 9). This is also seen when comparing the ability of these two double stranded oligonucleotides to compete the complex off control AP-1 probe. The 1514 AP-1 sequence does not compete the AP-1 complex off the control AP-1 sequence as well as excess cold control oligonucleotide (lane 11 compared to lane 12). This observation can be

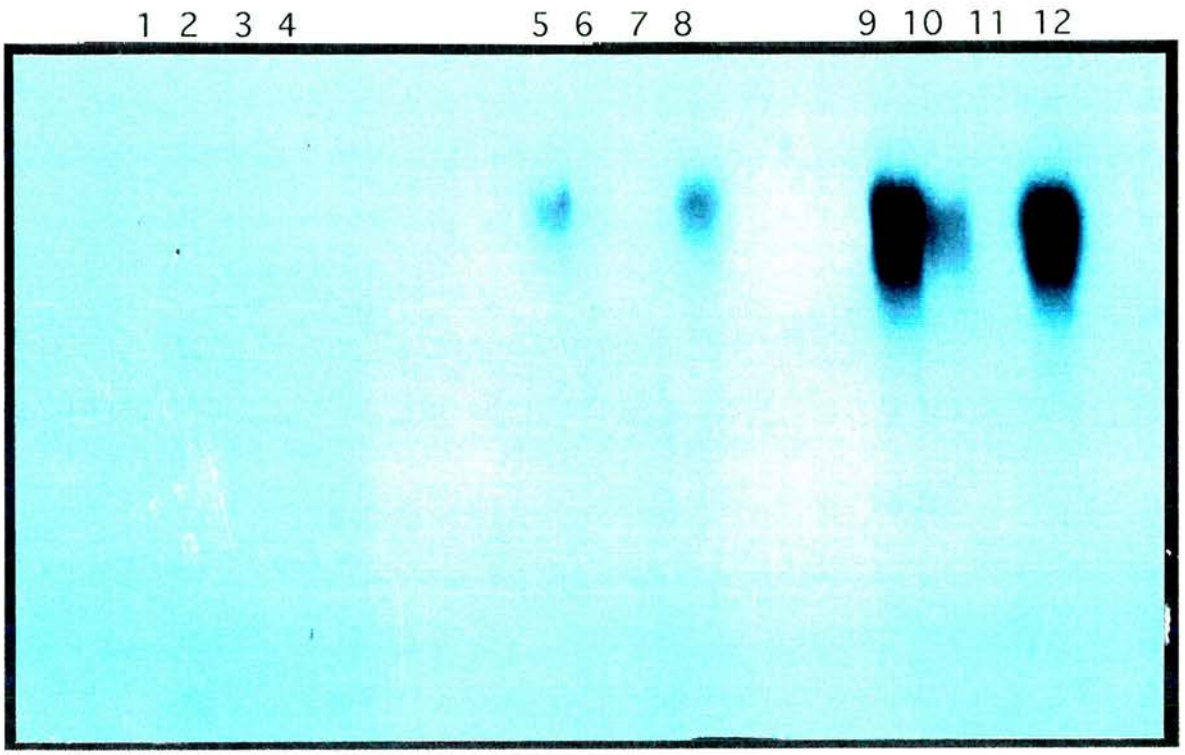
Figure 4.2

Comparison of 1514 and EV-1 TATA box proximal AP-1 sites. Fig.4.2a shows an autoradiograph from a standard (overnight) exposure, Fig 4.2b shows the same gel autoradiographed for 4 days. The control AP-1 tracks Fig.4.2a (9-12) are not shown in Fig.4.2b. Below the sequences of the 1514 and EV-1 oligonucleotides are shown.

1514 TGCTT**G**AGTCATAACC
EV-1 TGCTT**A**AGTCATAACC

These two sequences differ at a single base (bold) which lies within the AP-1 consensus site (underlined).

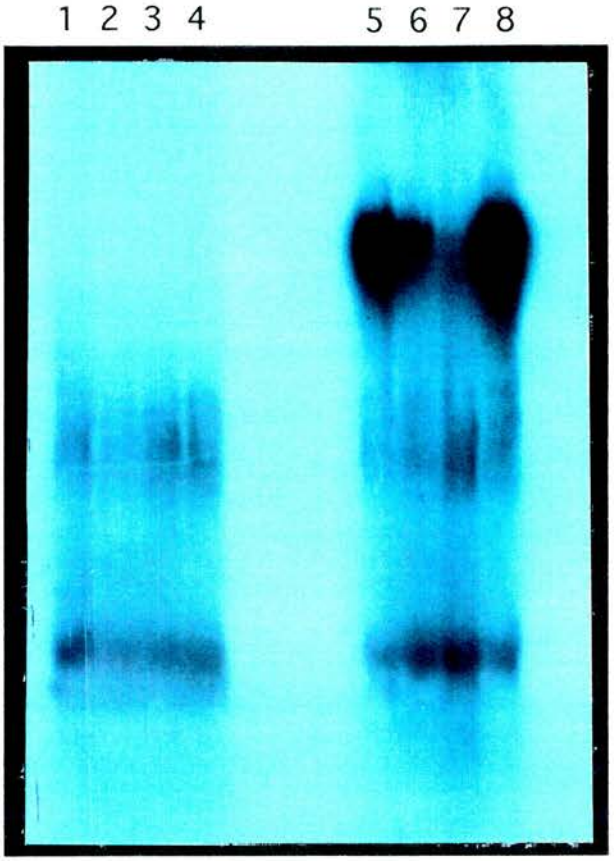
FIGURE 4.2a



PROBE - LANES 1-4 EV-1 AP-1
5-8 1514 AP-1
9-12 CONTROL AP-1

COMPETITORS - LANE 1/5/9 -
2/6/10 1514 AP-1
3/7/11 CONTROL AP-1
4/8/12 EV-1 AP-1

FIGURE 4.2b



PROBE - LANES 1-4 EV-1 AP-1
5-8 1514 AP-1
COMPETITOR - LANE 1/5 -
2/6 1514 AP-1
3/7 CONTROL AP-1
4/8 EV-1 AP-1

explained in two ways. Firstly two distinct complexes may be interacting with these two oligonucleotides. The cross competition seen between these two sequences suggests that this is not the case. The second possibilty is that the shorter flanking sequences around the core recognition site on the 1514 AP-1 sequence reduce the affinity of the AP-1 complex for this oligonucleotide when compared to the control AP-1 sequence. This second explanation is also supported by later observations with longer oligonucleotides (Section 4.3).

Even following longer autoradiography no binding to the EV-1 AP-1 sequence is observed (Fig.4.2b, lanes 1 to 4). This strongly suggests that AP-1 is not binding to this sequence. The observations made on the reduced AP-1 binding seen using the 1514 AP-1 sequence when compared to the control AP-1 sequence are also relevant here. While the results clearly show that the 1514 AP-1 sequence is capable of specific binding to a complex recognising the AP-1 sequence the binding may be limited by the size of the oligonucleotide. Thus, although the EV-1 AP-1 sequence may be capable of binding no interaction may be seen using this oligonucleotide due to the short flanking sequences. In order to ascertain if the absence of a specific shift on the EV-1 AP-1 sequence was due to the absence of sufficient flanking sequence, longer oligonucleotides were compared by gel shift analysis. From the data presented in this section it is not possible to state categoricaly that AP-1 is not interacting with the EV-1 sequence. However, it is clear that the EV-1 sequence is not functionaly identical to the 1514 sequence.

4.3 A second transcription factor site in close proximity to the AP-1 site

In the previous section, gel retardation experiments with oligonucleotides covering the AP-1 region of the EV-1 and 1514 LTRs showed distinct AP-1 binding patterns. This data could not be extended to exclude the possibility of AP-1 binding to the EV-1 AP-1 sequence because of potential problems caused by the size of the oligonucleotides.

In order to test more fully the ability of AP-1 to bind the EV-1 sequence, oligonucleotides with extended flanking sequences either side of the core AP-1 recognition site were obtained. Initially two oligonucleotides were prepared corresponding to the sequences of EV-1 LTR-1 and 1514. These two sequences differ at two positions, the mutation in the AP-1 site and also a second G to A transition. These two oligonucleotides and the results of gel retardation experiments using them as probes are shown in Fig.4.3.

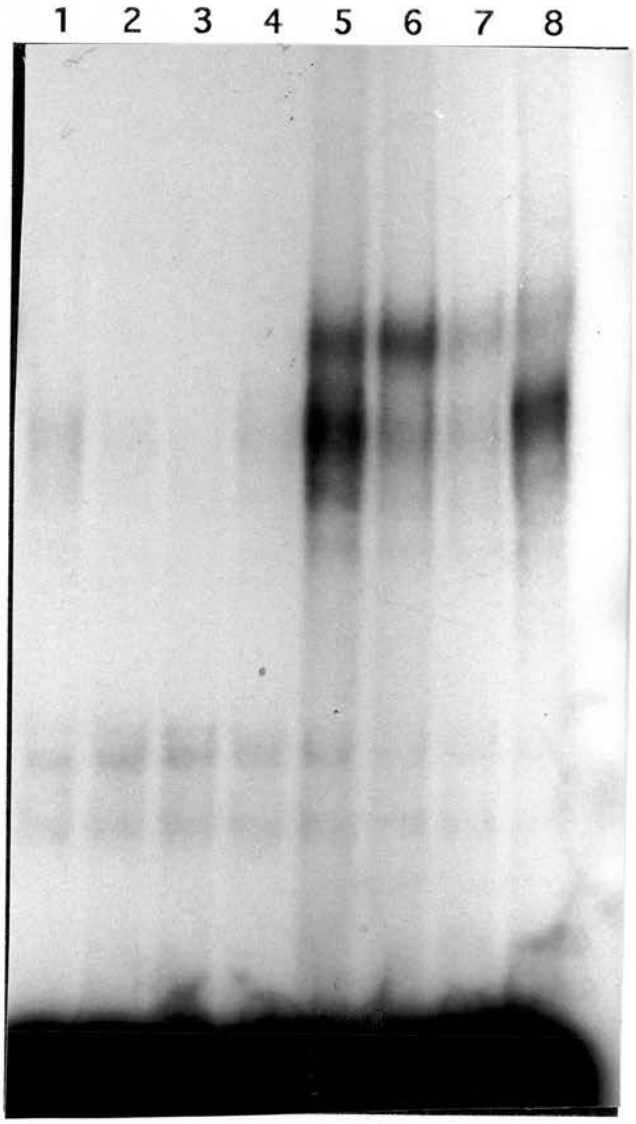
Figure 4.3

A gel shift using the oligonucleotides EV-1 *ex vitro* (EV-1) and 1514 is shown. Two specific retarded complexes are seen using the 1514 sequence (Lane 5 vs. 8). The EV-1 ex-v sequence shows a very weak shift corresponding to the lower band seen with the 1514 sequence. The two sequences are shown below.

1514 TGATGCTT**G**AGTCATAACC**G**CA
EV-1 TGATGCTT**A**AGTCATAACC**A**CA

In addition to the alteration in the AP-1 site these sequences also differ due to a second G to A conversion downstream (towards the TATA box) from the AP-1 site.

FIGURE 4.3



PROBE - LANES 1-4 EV-1 ex-vitro
" 5-8 1514

COMPETITORS - LANES 1/5 -
" 2/6 EV-1 ex-vitro
" 3/7 1514
" 4/7 AP-1

The results from Fig.4.3, where the EV-1 sequence was used as probe (lanes 1 to 4), show that even when the flanking sequence is extended there is no evidence of AP-1 binding. Weak bands were observed in lanes 1 and 4 with the EV-1 probe. The lengthened 1514 sequence generates, in addition to the AP-1 shift, a second retarded complex. The upper band is seen to be competed by both the 1514 (lane 7) and control AP-1 (lane 8) sequence confirming it as the same complex observed in Fig.4.3. The control AP-1 sequence, unlike the 1514 oligo, cannot compete the lower band. Suggesting that the lower complex is recognising a sequence distinct from the AP-1 consensus site. This second, non-AP-1 complex, is competed by the EV-1 sequence even though this oligonucleotide cannot compete the complex recognising the AP-1 site (lane 6). The weak bands seen in lanes 1 and 4 where the EV-1 sequence was used as the probe suggest that the EV-1 sequence has a low affinity binding site for this factor.

The EV-1 and 1514 oligonucleotides described above differ at two positions, one in the AP-1 region the other outside the AP-1 site (Fig.4.3). As the second factor does not appear to be recognising the AP-1 site it appeared likely that this G to A conversion outside the AP-1 recognition site was responsible for the reduced affinity of the EV-1 oligonucleotide for this sequence. In the EV-1 LTRs the second, non-AP-1 mutation, is not found in all LTR sequences. In the *ex-vivo* LTRs, and a subpopulation of the *ex-vitro* LTRs, the 1514 sequence is found.

In order to test whether this second factor was binding differently to the two EV-1 sequences, the corresponding region of the *ex-vivo* EV-1 LTRs (EV-1ex-v) was compared to the EV-1 and 1514 sequences in gel retardation experiments Fig.4.4. Using the EV-1ex oligonucleotide as a probe did not generate an AP-1 specific shift (lanes 11 to 15), confirming the results with the EV-1 sequence and demonstrating that the base change outside the core AP-1 recognition site does not affect AP-1 binding. While the EV-1 ex-v sequence does not bind an AP-1 complex it does bind to the second factor with an affinity apparently equivalent to that for the 1514 sequence (lane 15 compared to lane 10). This confirms that the recognition site for this factor is distinct from the AP-1 recognition site and lies in the 3' flanking region of the oligonucleotide.

In order to test this more fully, an oligonucleotide covering the 3' region of the 1514 and EV-1ex-v oligonucleotides was generated. This sequence was designated VSS (Visna specific sequence). When this oligonucleotide is used as a probe (Fig.4.5) it generates a

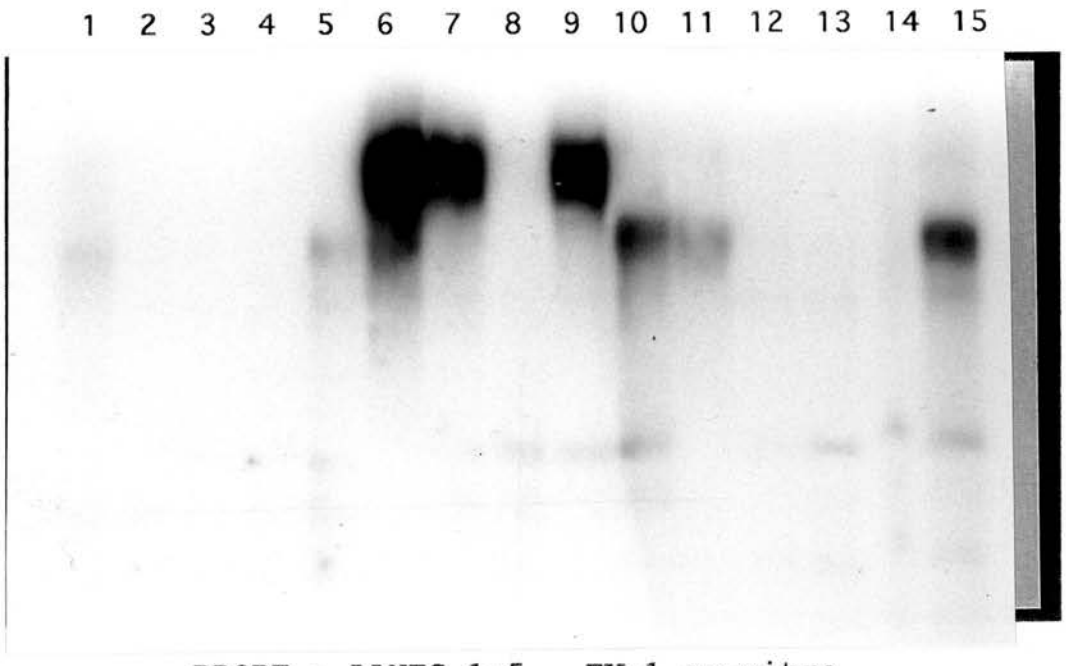
Figure 4.4

The three oligonucleotides EV-1, 1514 and EV-1 ex-v (the sequence from the *ex vivo* EV-1 LTRs) are compared in a gel shift assay. The bands for the EV-1 and 1514 sequences are as in Fig.4.3. The EV-1 ex-v sequence shows a single shifted band corresponding to the lower band seen with the 1514 oligonucleotide. This band is of higher intensity with the EV-1 ex-v than the EV-1 sequence (Lane 1 vs 11). The three sequences are shown below.

1514	TGATGCTT**G**AGTCATAACC**G**CA
EV-1ex-v	TGATGCTT**A**AGTCATAACC**G**CA
EV-1	TGATGCTT**A**AGTCATAACC**A**CA

The varying bases are again shown in bold. The EV-1 sequence differs from both the EV-1ex-v and 1514 due to the G to A conversion outside the AP-1 site.

FIGURE 4.4



PROBE - LANES 1-5 EV-1 ex-vitro
" 6-10 1514
" 11-15 EV-1 ex-vivo

COMPETITORS - LANES 1/6/11 -
" 2/7/12 EV-1 ex-vitro
" 3/8/13 1514
" 4/9/14 EV-1 ex-vivo
" 5/10/15 AP-1

Figure 4.5

The oligonucleotide VSS is generates a single non-AP-1 shift (Lanes 1 and 4). The additional band observed in lanes 2,3 and 5 is non-specific. The sequence of the VSS oligonucleotide is shown below.

```
1514  TGATGCTTGAGTCATAACCGCA
VSS             GAGTCATAACCGCA
```

The VSS oligonucleotide corresponds to the 3' half of the 1514 sequence and does not contain the AP-1 site.

FIGURE 4.5

PROBE - VSS

COMPETITORS LANE 1 -
2 VSS
3 1514
4 AP-1
5 OCT-1

sequence specific retarded complex (lane 1). The complex binding to this oligonucleotide is not competed by excess oligonucleotide containing only the AP-1 site (lane 4). This, together with the lack of a consensus AP-1 sequence on the VSS oligonucleotide, provides strong evidence that this shift is not related to the complex binding the AP-1 site. The OCT-1 sequence (lane 5) was used as a competitor as comparison of the oligonucleotide sequence with known consensus sequences (Faisst & Meyer 1992) suggested that the recognition site bore closest similarity to the OCT recognition site.

Using this VSS oligonucleotide as a competitor, Fig.4.6, it can be seen this sequence will compete for the unknown complex without interfering with the AP-1 complex (lane 5). This gel also illustrates the independence of these two factors in relation to their ability to bind DNA. Both can be competed without affecting the binding of the other. This appears to rule out synergistic binding when both sites are present, as in the 1514 LTR. A second point is whether both factors are capable of binding to this region simultaneously.

Binding of both factors to the 1514 region should result in the generation of a third retarded complex migrating above the shifts caused by the single protein complexes. In no experiment using the 1514 sequence as a probe was this third complex seen. This may reflect the fact that in these experiments the probe was not saturated. This may have prevented the formation of the two protein complex. Experiments were performed to attempt to saturate the 1514 probe using reduced probe concentrations and varying the quantity of nuclear extract used. Despite varying NaCl concentration over the range of 50-130mM it was not possible to saturate the probe. However, up to 50% of the probe was present in the retarded bands (as determined by liquid scintillation counting). This would give an expected approximatly 5% of probe in the double complex, assuming no interference between the factors. A band of this intensity should have been readily visible on the autoradiographs but was never observed.

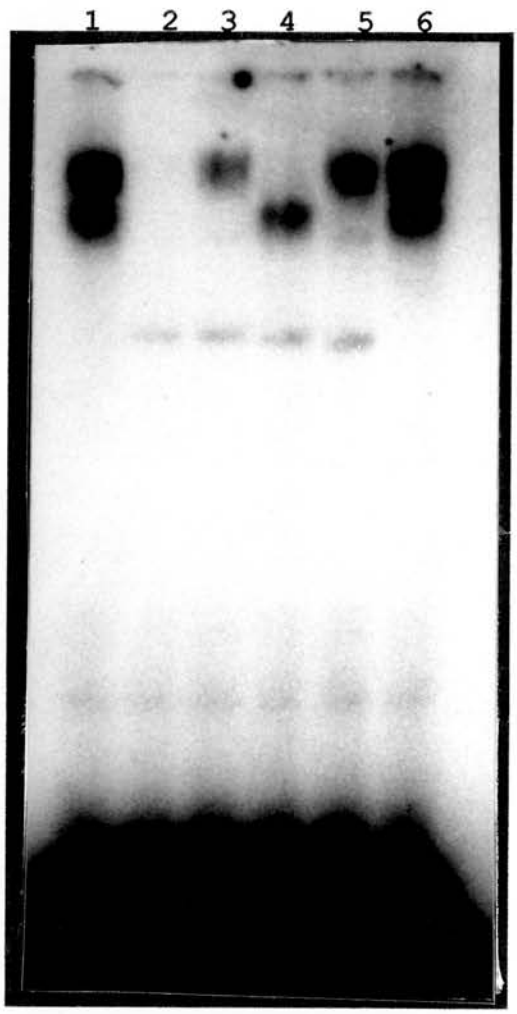
4.4 The upstream region of the MVV LTR contains a second copy of the VSS Sequence

Many promoters both viral and cellular contain multiple recognition sites for the same factor. In light of this the LTR was checked for any additional sequences with homolgy to the VSS sequence. This search revealed a possible additional site for this factor, position 85 in both the 1514 and EV-1 LTRs (Fig3.4). This sequence underlies the footprint observed by

FIGURE 4.6

The AP-1 site and the VSS sequence are independent sites within the 1514 LTR. Using the 1514 oligonucleotide as a probe it can be seen that the AP-1 and VSS bands may be independently competed (lanes 3 and 4).

FIGURE 4.6



PROBE - 1514

COMPETITOR - LANE 1 -
2 1514
3 EV-1 ex-vivo
4 AP-1
5 VSS
6 OCT-1

Gabuzda *et al.* (1989). This region of the EV-1 and 1514 LTRs is compared in Fig.4.7.

In addition to a potential site for the unknown factor this region also contains a degenerate AP-1 site and a consensus E-box site (CANNTG). An oligonucleotide covering this region from the 1514 LTR was obtained and designated the USR (UpStream Region) oligonucleotide (Fig.4.8). When used in gel retardation experiments this sequence showed a single retarded band. This retarded band has the same sequence recognition as the protein responsible for the shift seen with the VSS oligonucleotide (lane 3). Specific competition of this protein with the VSS oligonucleotide does not reveal any additional transcription factor binding sites in this region of the LTR. This would suggest the AP-1 site in this region is non functional. The presence of E-box binding proteins was not assayed in the extracts used for these experiments, so it cannot be concluded that this site is non-functional in the MVV LTR. The conservation of the USR sequence between the 1514 and EV-1 LTRs shows the importance of this region in both sets of LTRs and points to its functionality in the EV-1 sequences.

The presence of two sites for this factor in the MVV LTR may be important for the regulation of transcription by the LTR. It also raises questions regarding differential regulation of the 1514 and EV-1 viruses. The structure of DNA at the promoter is not believed to be linear; rather the binding of transcription factors results in bending of the DNA. This effect allows transcription factors to interact directly even if they are separated by long sequences. Transcription factors such as Sp1 can directly mediate DNA bending (Su *et al.*, 1991). One possibility is that the two sites for the non-AP-1 factor in the MVV LTR are involved in the generation of higher order promoter structures. As there appears to be a difference in the transcription factors binding to the 1514 and EV-1 LTRs it is interesting to speculate how distinct the mechanisms of transcriptional regulation are in these two viruses. This will be discussed in greater detail in Section 5.5. Having identified a novel transcription factor site in the 1514 and EV-1 LTRs it was decided to examine the factor binding to the VSS region in more detail.

FIGURE 4.7 Putative Upstream VSS

```
                60      AP-1  E-BOX                    97
1514            TCAGGATGACACAGCAAATGTAACCGCAAGTTCTGCTT
LTR 1,28        ----------T----------------GCGCAAAG
LTR 19          ----------.----------T------GCGCAAAG
LTR30           ----------------------------GCGCAAAG
LTR18,48        ----------------------------.---C----
LTR44           ---------------------------------C----

VSS                             GAGTCATAACCGCA
```

Sequences from the various LTRs are aligned with the upstream region footprinted in the 1514 LTR. The region which footprints in the 1514 LTR is underlined. This region also contains putative AP-1 and E-Box sites. In EV-1 LTRs 1,19,28 and 30 the 90 to 97 is no longer conserved as these sequences lack the position 90 insert. The sequence of the VSS oligonucleotide is shown below these upstream sequences and the sequence conserved compared to the 1514 footprint is underlined.

FIGURE 4.8

The sequence of the USR oligonucleotide is shown above the autoradiograph. Putative AP-1, E-Box and VSS sites are shown. From the data it can be seen that only the VSS generates a specific shift.

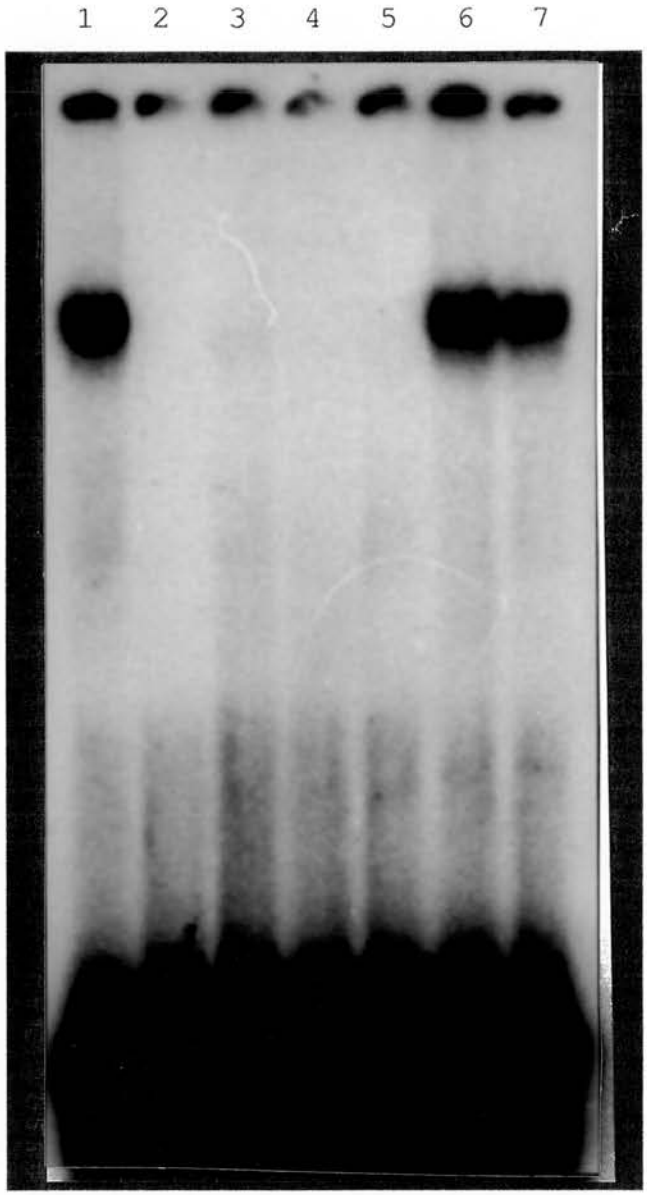
FIGURE 4.8

USR sequence

-120 -90

AP-1? E-BOX

TCAGGATGACACAGCAAATGTAACCGCAAGTTCTGCTT



PROBE - USR

COMPETITORS LANE 1 -
2 USR
3 VSS
4 EV-1ex-vivo
5 1514
6 AP-1
7 OCT-1

4.5 The VSS Binding Protein also Interacts with the TCR Alpha Gene Enhancer

The sequence boundary of the non-AP-1 factor in the MVV LTRs is not listed as being recognised by any characterised transcription factor (Faisst & Meyer 1992). While there has been no characterisation of the VSS binding protein, a closely related sequence has been footprinted within the TCR alpha gene enhancer (Ho *et al.*, 1989 & 1990). This footprinted region, over box 4 (Tα4) in the enhancer, when aligned with the sequence present in the MVV LTRs (Fig.4.9) shows a conserved sequence. The Tα4 sequence was used in gel shift experiments in comparison with the MVV sequences (Fig.4.10a/b). Cross competition shows that the binding specificity of these sequences is the same despite the divergence between the TCR enhancer and MVV sequences. Alignment of these sequences which bind this factor gives a consensus binding site for this unknown protein of $^{T}/_{C}$AACC$^{A}/_{G}$CA. The divergence from this sequence in LTR 1 at the TATA box proximal site, within the VSS sequence, appears to result in reduced binding activity following the G to A conversion at position six. Having identified the recognition sequence of this factor it was decided to look at other characteristics of the protein. Studies on the TCRα gene enhancer revealed that the factor binding box 4 required other sequences within the enhancer to function (boxes 1 and 2) and deletion only had a minor effect on enhancer activity (Ho *et al.*, 1989). This data suggest that the VSS sequence is not capable of functioning as an enhancer element and so must be positioned at the promoter either indirectly, by proximity to a factor which can remodel chromatin structure and bring it into contact with the promoter element or by being present within the promoter sequence itself. The possible function of this sequence in relation to a basal promoter has not been tested, only its inability to function as an enhancer element has been shown. This question of whether the VSS sequence can function as a transcriptional activator is addressed in Chapter 5.

As it appears that in the MVV LTR the same factor is binding the upstream and downstream sites the footprinted regions for these two sites were compared (Fig.4.11). Comparison of the area protected from DNase footprinting suggests that at the down stream site the VSS site is not occupied. This provides some additional indirect evidence to suggest that these sites may not be occupied simultaneously and that these two factors AP-1 and the VSS binding protein, are competing to bind to this region of the 1514 LTR.

FIGURE 4.9 Comparison of Tα4 with the MVV VSS

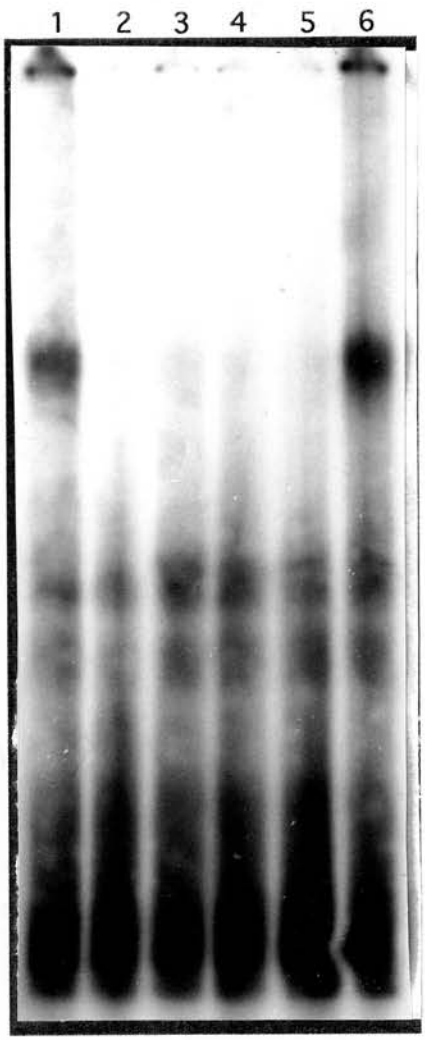
Tα4 Footprint	CCCCCAACCGCAGGTGC
1514 Footprint (USR)	AAATGTAACCGCAAGTTCT
1514 TATA box proximal	TGAGTCATAACCGCAGATGT
EV-1 TATA box proximal (LTR 1)	TAAGTCATAACCACAATTGT
VSS oligo	GAGTCATAACCGCA

The underlined sequence represents a putative core recognition site for the factor binding this sequence. From this comparison the core recognition sequence appears to be non-palindromic. The sequences denoted as 'footprint' are those regions which are protected from DNAse I digestion

FIGURE 4.10

Comparison of the oligonucleotides VSS and Tα4 by gel shift. Fig.4.10a shows that when the VSS oligonucleotide is used as a probe the specific band is competed by both the MVV LTR sequences and the Tα4 sequence (Lane 2). The reciprocal experiment is shown in Fig.4.10b.

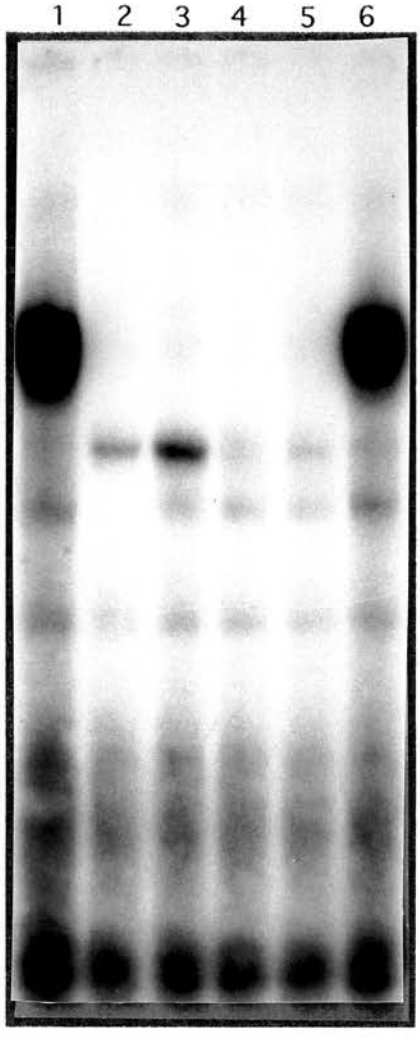
FIGURE 4.10a



PROBE - VSS

COMPETITOR - LANE 1 -
" 2 Talpha 4
" 3 VSS
" 4 1514
" 5 USR
" 6 Control AP-1

FIGURE 4.10b



PROBE - Talpha 4

COMPETITOR - LANE 1 -
" 2 Talpha 4
" 3 VSS
" 4 1514
" 5 USR
" 6 Control AP-1

FIGURE 4.11 The TATA box proximal VSS/AP-1 element in the 1514 LTR

The 1514 oligo covering this region binds these two factors in an apparently independent manner. This would suggest that only one or other of these two sites may be occupied. Below the footprinted region is highlighted in bold type and the VSS and AP-1 consensus, 'core', sequences are shown.

	AP-1
Oligo sequence	TG**ATGCTTGAGTCATAACCGC**A
	VSS
Upstream region	GC**AAATGTAACCGCAAGTTCT**GC

This data on footprints is taken from the paper by Gabuzda et al (1989). In both cases the two footprints are assumed to be caused by single factors. The gel retardation data using the USR oligonucleotide suggests that the footprint in this region is caused by the same factor as binds in the TATA box proximal region. Taking these two assumptions this then several conclusions follow. Firstly that the VSS protein was saturating in the system (the USR footprint). Secondly, in the TATA box proximal region the footprint does not cover the whole VSS 'core'. Thirdly the VSS factor protects regions outside its 'core' sequence (upstream footprint).

From these observations together with the experimental data presented previously in the gel shift experiments it appears unlikely that these two factors are binding simultaneously.

4.6 Partial Characterisation of the VSS binding protein

4.6.1 The VSS binding factor is antigenically distinct from the complex binding the AP-1 site

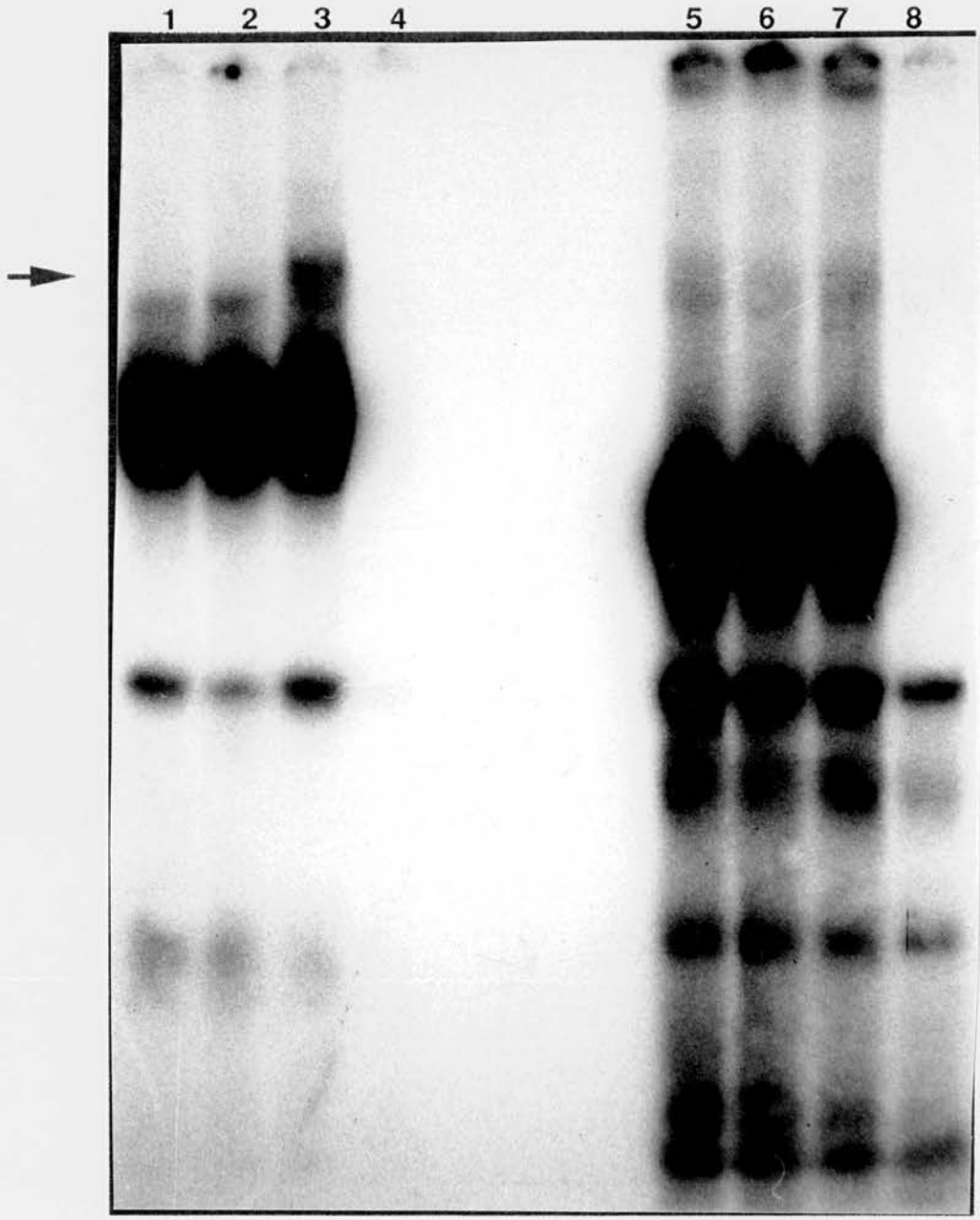
The relationship between the factors binding to the VSS and AP-1 sequences in the LTR has so far been described solely in terms of the sequence recognition of these two factors, which appear distinct and unrelated. By cross competition experiments in gel shift assays it is clear that the two transcription factors recognise distinct sequences. However, in both of the sequences tested from the MVV LTR the VSS sequence appears near a conserved (1514, TATA box proximal site) or degenerate (EV-1 TATA box proximal site and USR site in 1514 and EV-1) AP-1 site. The Jun family of proteins are known to interact with transcription factors other than those produced by the *c-fos* gene family. One well characterised example is the NF-AT complex generated in activated T cells (reviewed by Rao 1994). The interaction of Jun with the EV-1 LTRs, in the absence of an AP-1 site, would allow the same general mechanism of Tat transactivation and LTR induction in the EV-1 and 1514 viruses. If Jun is not interacting with the EV-1 LTRs at any position then it is possible that this virus is using a distinct system to control its induction and possibly Tat responsiveness.

In order to test for the presence of the Jun protein in the VSS binding complex an anti-Jun polyclonal antibody was used in gel supershift experiments. This antibody binds all identified members of the *c-jun* gene family; Jun, Jun B and Jun D. The antibody was raised against a peptide from the DNA binding domain of the mouse Jun protein and is cross reactive with the human, chicken and rat proteins. When antibody is added to the gel retardation reactions prior to loading onto the gel it can stably bind protein/DNA complexes. This generates a band which migrates above the normal complex. This technique can therefore be used to identify components of the retarded complex. In Fig.4.12 the anti-Jun antibody generates a supershift on the 1514 AP-1 oligonucleotide but no supershift when the VSS oligonucleotide is used as a probe. This data does not exclude the possibility of a Jun family protein being present in the VSS binding complex but does illustrate that if present it is not in the same conformation as in AP-1. These data together with the work on identifiying the core region of the VSS strongly suggests that this factor is distinct from AP-1. It also demonstrates that this reagent is cross reactive with the ovine Jun proteins.

FIGURE 4.12

The 1514 AP-1 and VSS oligonucleotides were used as probes. The specific retarded band generated by the 1514 oligonucleotide can be seen to contain a Jun family protein as demonstrated by the specific supershift, obtained after incubation with a rabbit anti-Jun antisera, (Lane 4) which is marked by an arrow. No supershifted band is observed with the VSS specific complex (Lane 7).

FIGURE 4.12



PROBE - LANES 1-4 1514 AP-1
- " " 5-6 VSS

ANTISERUM - LANES 1/5 -
" " 2/6 Control IgG
" " 3/7 anti-Jun IgG

LANE 4 - 1000x AP-1
LANE 8 - 1000x VSS

4.6.2 UV crosslinking to the VSS sequence

A technique which has been used in several systems for the identification of proteins binding to an oligonucleotide is to covalently link the protein to the DNA via a UV excitable group incorporated into the recognition sequence. As three distinct sequences which bind the non-AP-1 complex have been identified it was possible to generate an oligonucleotide with a BrdU substitution in the putative recognition site.

The cross linking assay was performed as described in Chapter 2. Complexes were formed as in the gel retardation experiments (Section 4.1) with the exception being the use of the BrdU substituted oligonucleotide being used as the probe. These reactions were placed on ice, to avoid evaporation and over heating of the sample, and placed in a commercial UV cross linker for the period of time indicated in each experiment.

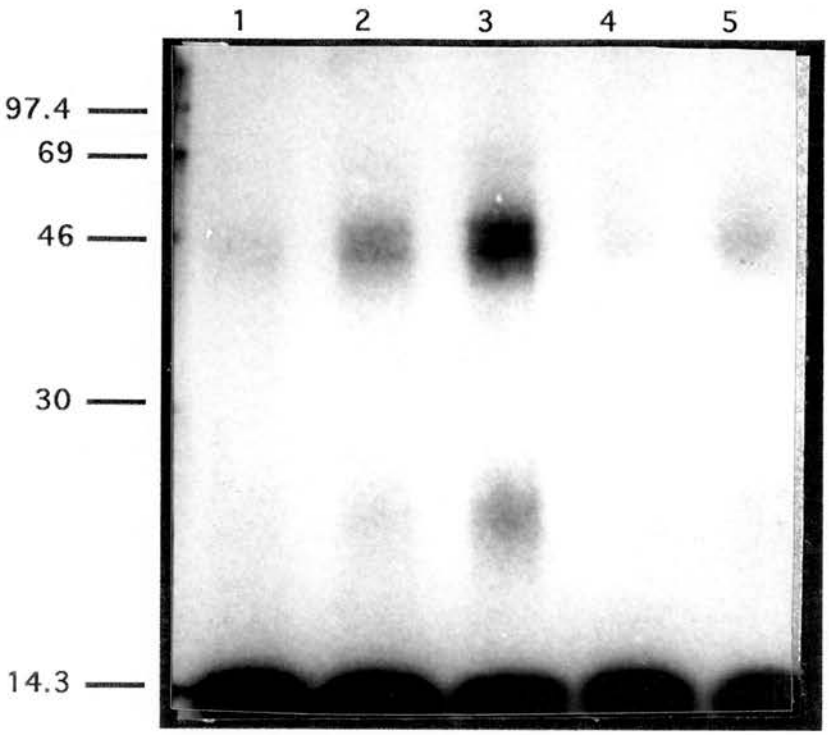
The first step in this process was the calculation of the optimum exposure to UV required for the generation of a detectable crosslinked complex. The results of a representative time course are shown in Fig.4.13. This Figure shows a gradual accumulation of the covalently linked retarded complex as the length of exposure to UV is extended. From this experiment an exposure time of 1 hour was used. In lanes 4 and 5, 1000x fold cold VSS oligonucleotide was present. The comparison between the bands seen in these lanes and those where no competitor was present suggests the band seen at 40-46kD is specific for the BrdU substituted VSS oligonucleotide. This 40-46kD band appears to contain two distinct species with closely related molecular weights.

Non-specific competitor may interfere with specific complexes if it is present at excessive concentrations. The level of poly dI-dC.dI-dC used in standard gel retardation experiments was 5μg per reaction. This level was determined empirically for the standard gel retardation assay and may not be optimum for the cross linking assay. In order to determine whether the ideal level of non-specific competitor was different for these assays in comparison to the standard gel retardation protocol, a titration of poly dI-dC.dI-dC was performed (Fig.4.14). In the absence of competitor four bands appear. As the level of non-specific competitor increases all the complexes other than that at 40-46kD are strongly competed. The 40-46kD complex is only reduced at high levels of non-specific competitor suggesting it represents the only specific complex formed on the BrdU VSS oligonucleotide. From these experiments it appears that the 5μg of poly dI-dC.dI-dC used in the gel retardation

FIGURE 4.13

Cross linking timecourse using the BrdU substituted VSS oligonucleotide as a probe. The figures on the left denote molecular weight in kDa. The strong band at the bottom of the gel is the probe migrating with the dye front.

FIGURE 4.13



PROBE - X-LINK VSS

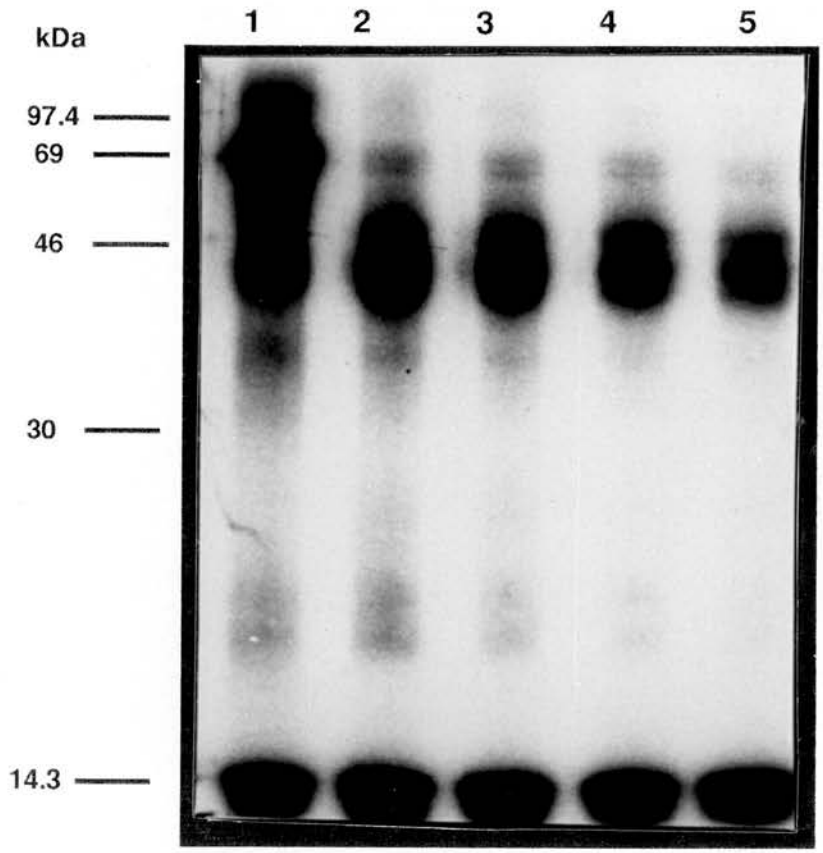
COMPETITOR - LANE 1/2/3 -
" 4/5 VSS

UV EXPOSURE - LANE 1 10 minutes
" 2/4 30 minutes
" 3/5 60 minutes

FIGURE 4.14

Titration of the non-specific competitor poly dI.dC-dI.dC. for the UV crosslinking experiments. Non-specific competitor was titrated over the range 1-10μg.

FIGURE 4.14



PROBE - X-LINK VSS

COMPETITOR - LANE 1 -
" 2 1μg dIdC
" 3 3μg dIdC
" 4 5μg dIdC
" 5 10μg dIdC

experiments is also suitable for this protocol. Levels below this generate quite high backgrounds and at 10μg there is interference with specific bands.

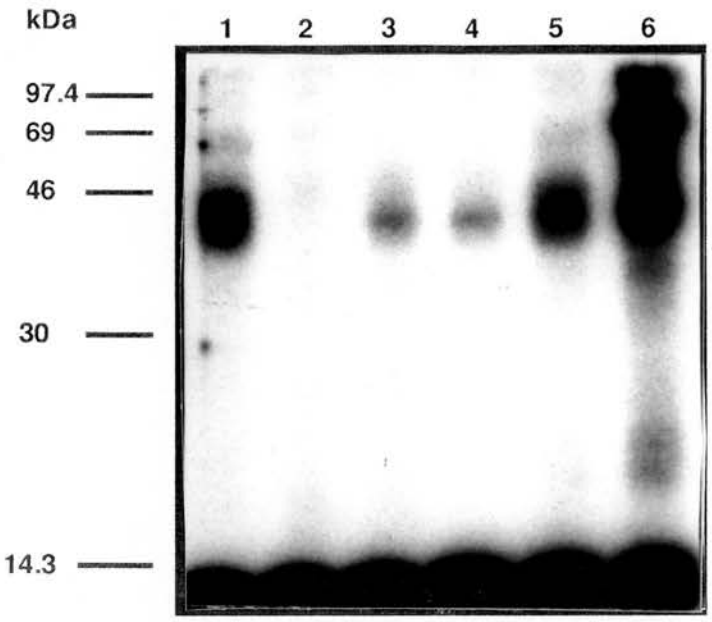
Having calculated the optimal incubation time and non-specific competitor levels experiments using specific competitors were performed. The results of the competition experiments are shown in Fig.4.15. In the presence of specific competitor, only the 40-46kD complex is competed. All VSS containing sequences show competition against this complex but no effect is seen with the oligonucleotide containing the AP-1 sequence. The incomplete competition levels seen with the VSS and 1514 oligonucleotides, versus the USR sequence, is readily explained by the observation that in the VSS and 1514 sequences the binding site is at the extreme end of the oligonucleotide. The final base pair of the putative binding site is the last base pair of the oligonucleotide. In contrast the USR sequence contains long flanking sequences either side of the recognition sequence. This is likely to reduce the affinity of binding and increase the on off rate of the factor on the VSS and 1514 sequences in comparison with the USR sequence. In the gel shift assay this effect was less noticable as the system was in equilibrium. In the cross linking experiments any complex forming with the VSS probe has a possibilty of becoming covalently linked thus even a slight reduction of competition efficiency will result in the presence of a detectable band and incomplete competition.

The complex seen at 40-46kD in the cross linking experiments corresponds to a protein of approximate MWT of 34-40kD after subtraction of the oligonucleotide MWT. The suggestion that this complex consists of two distinct bands may reflect some heterogeniety in the composition of the VSS binding complex.

FIGURE 4.15

Use of specific competitors in UV crosslinking. The specific competitors listed were used to demonstrate a specific interaction between the 36kDa band and the BrdU substituted VSS probe.

FIGURE 4.15



PROBE - X-LINK VSS

COMPETITOR - LANE 1 -
" 2 USR
" 3 VSS
" 4 1514
" 5 CONTROL AP-1
" 6 NO dIdC.dIdC

4.7 The VSS binding activity and AP-1 are Differentially Induced by PMA or TNFα

The results discussed in sections 4.2 & 3 suggest that the transcription factor AP-1 does not interact with the EV-1 LTRs at the TATA box proximal AP-1 site. One possibilty is that the second factor which binds in this region may be compensating for the lack of AP-1 binding in this region of the EV-1 LTRs. As the AP-1 complex interacting with the downstream site in 1514 has been described as the major regulator of inducible viral gene expression (Hess *et al.*, 1989) it is possible that this second factor was playing a similar role in EV-1. The presence of two sites for this factor in the LTR suggest it may play a major role in the induction of viral transcription following cellular activation. This hypothesis was initially tested by examining the response of the VSS binding protein to cellular mitogens.

The transcription factor AP-1 is induced by a large number of stimuli including cellular activation by cytokines and mitogens (reviewed by Herschman 1991, Nathans & Ryder 1988). The induction potential of the VSS binding protein was assayed in comparison to AP-1 activity in response to two distinct activation stimuli, PMA and ovine TNFα. Both of these stimuli are known to elevate the activity of AP-1 (Hanazawa *et al.*, 1993, Hass *et al.*, 1991). For these experiments cells were grown in serum free media for 24 hrs in order to stop cell growth and reduce levels of inducible factors to an absolute minimum. The half lives of the various Fos and Jun proteins is reported to range between one and four hours (Carillo S. *et al.*, 1994, Gruda *et al.*, 1994). During this incubation there were no obvious gross alterations in cell morphology or viabilty. This is important as removal of growth factors has been reported to cause apoptosis and the activation of the *c-fos* and *c-jun* genes in some cell systems (Colotta *et al.*, 1992). This effect was not observed in the chondrocytes used in these studies. Chondrocytes appeared able to tolerate over 48 hrs in serum free media without serious loss of viability. The level of AP-1 in extracts did not appear to be greatly affected by extending the serum withdrawal from 24 to 48 hours. The 24hr period was chosen as it should provide a suitable length of time to reduce inducible transcription factor levels without stressing the cells.

After this 24hr incubation the media on the cells was supplemented with serum free media containing the relevant mitogen. The concentration of mitogens used was 100ng/ml for both PMA and recombinant ovine TNFα (Green *et al.*, 1993, provided by B. Ebrahimi).

The results of these stimulation experiments are shown in Fig.4.16 and Fig.4.17, for PMA and TNFα respectively. For these experiments equivalent quantities of protein were used from each nuclear extract. Stimulation of AP-1 binding activity is seen following both stimuli. The main problem encountered during these experiments was the level of AP-1 activity seen in unstimulated cells. This would suggest that the chondrocytes used in this study express the *jun* genes, and probably the *fos* genes, at relatively high levels. From these studies it was not possible to describe which members of the *jun* and *fos* families were induced by these stimuli. From Fig.4.16 it can be seen that 5% serum, as expected, also elevates AP-1.

While the AP-1 complex shows the expected induction with these mitogens there was no consistent evidence for an induction of the VSS binding protein. This provides a second piece of indirect evidence against the VSS binding activity being related to AP-1 or its constituent proteins. VSS levels remain constant despite stimulation of cells with PMA, ovine TNFα or 5% serum. While the VSS binding factor appears to be constitutive, and uninduced by these stimuli, this does not mean that it is unable to mediate inducible transcription. This could occur either by a modification of the DNA binding complex itself or the recruitment of inducible factors by this protein once it has bound DNA. So it appears that the role of this protein in LTR induction, if it has one, is at best secondary. A further point is that the chondrocytes may not be in resting state at the t=0 time point, as seen by the strong AP-1 activity found in the cells. It is therefore possible that the factor binding the VSS is already maximaly induced in this cell type. One system which should be examined for induction of this would be monocyte to macrophage differentiation, as this appears to be the trigger *in vivo* for viral replication.

4.8 Discussion

In the *in vitro* gel retardation assays described in this chapter no evidence was found for the interaction of AP-1 with the degenerate TATA box proximal site in the EV-1 LTRs. This site has been described as the main target for the regulation of inducible viral gene expression in the 1514 virus (Hess *et al.*, 1989). It has also been claimed that in 1514 Tat transactivation is absolutely dependent on the expression of *c-jun* (Neuveut *et al.*, 1993). The failure to observe AP-1 binding does not appear to be due to a limitation of the protocol used here as

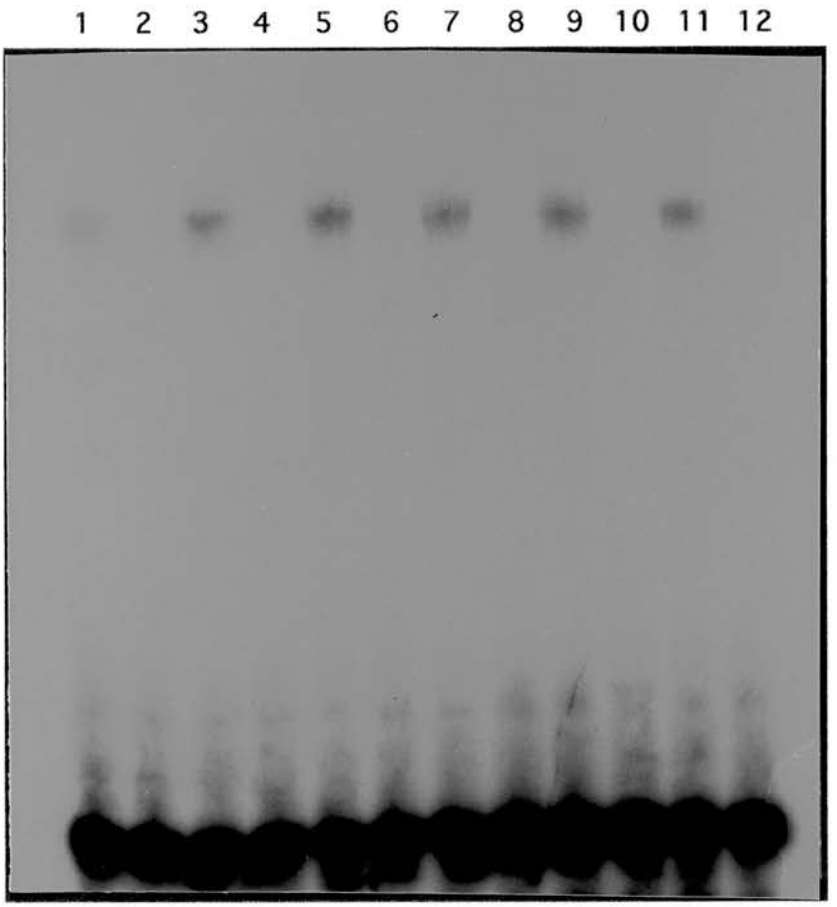
FIGURE 4.16

Timecourse for PMA stimulation. The induction of the VSS binding protein and AP-1 by PMA and serum was examined. The AP-1 probe is used in the top panel and the VSS probe the bottom panel.

Lanes 1/2 - Unstimulated
Lanes 3/4 - 1 hours post PMA
Lanes 5/6 - 2 hours
Lanes 7/8 - 4 hours
Lanes 9/10 - 6 hours
Lanes 11/12 - 5% FCS (12hours)

Even numbered lanes contain 1000x specific competitor. For this experiment induction was quantitated by cutting out the retarded band and counting by liquid scintillation. The values below are for three independent experiments.

	AP-1	VSS	
1 hr	1.8x	1.1x	(fold induction)
2hrs	2.2x	1.3x	
4hrs	1.7x	1.2x	
6hrs	1.5x	1.1x	
5%FCS	1.7x	0.9x	


1 2 3 4 5 6 7 8 9 10 11 12

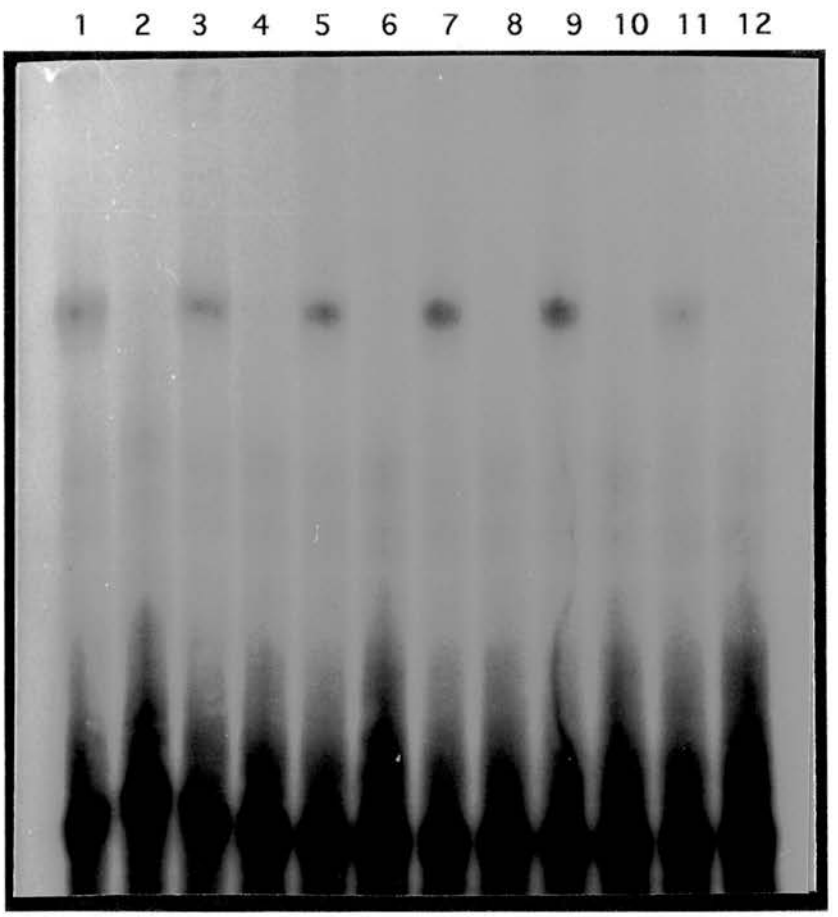

1 2 3 4 5 6 7 8 9 10 11 12

FIGURE 4.17

Induction of AP-1 and the VSS binding protein by ovine TNFα. Chondrocytes were stimulated with 100ng/ml TNFα. The top panel is for the AP-1 probe and the bottom panel for the VSS probe.

Lane 1/2 - Unstimulated

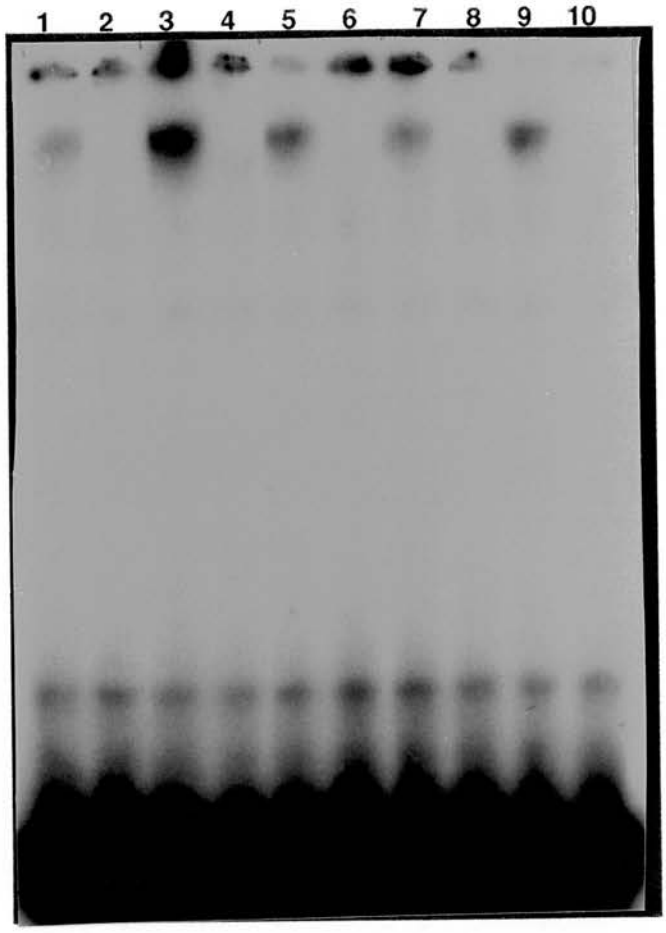
Lane 3/4 - 1 hour post TNFα

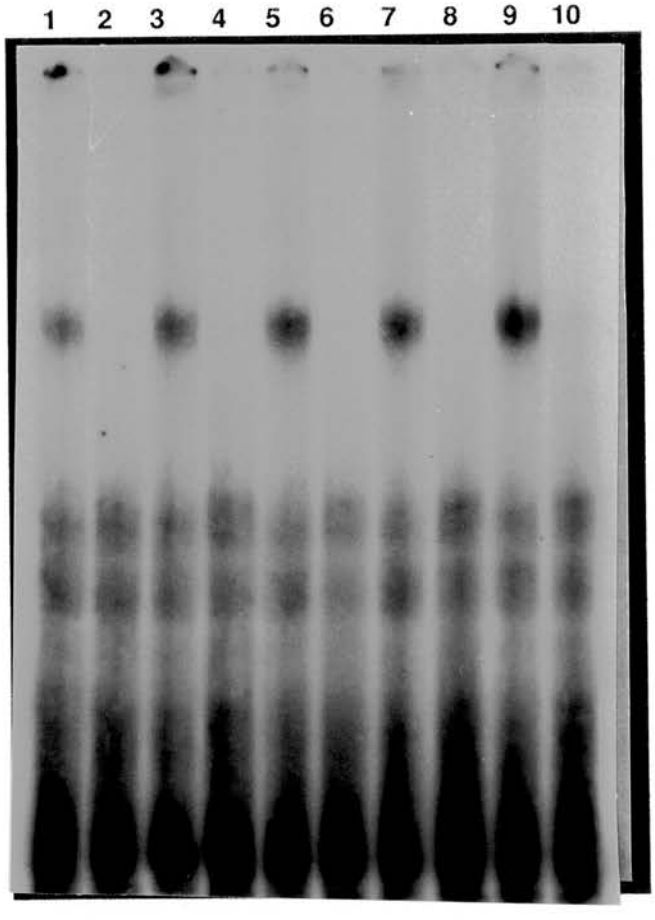
Lane 5/6 - 2 hours

Lane 7/8 - 4 hours

Lane 9/10 - 6 hours

Even numbered lanes contain 1000x excess specific competitor.


1 2 3 4 5 6 7 8 9 10


1 2 3 4 5 6 7 8 9 10

strong AP-1 binding to the 1514 TATA box proximal site was observed. As was discussed in Chapter 3, this TATA box proximal site represents the best match to the consensus AP-1 sequence in the EV-1 LTR (Fig.3.7). As AP-1 has been described a major regulator of MVV LTR activity how does the EV-1 virus compensate? While there is no evidence of AP-1 binding to the TATA box site it is possible that one of the other sites is indeed functional. DNase footprinting experiments (Gabuzda *et al.*, 1989) and comparison of the degenerate sequences within the LTR (Section 3.4) with the AP-1 consensus site argues strongly against this.

Leaving aside the possibility of one of the degenerate AP-1 sites being functional then two distinct models could be invoked. Firstly the Jun protein may interact with the LTR, at a site distinct from the TATA box proximal site, as a subunit of a transcription factor recognising a non-AP-1 site. This would still be consistent with the requirement of *c-jun* expression for Tat action. The second model would involve EV-1 Tat transactivating through distinct *cis*-acting sequences in the LTR. This would remove the requirement for AP-1 or Jun interaction with the LTR. As the functional domains of MVV Tat have only been partially characterised (Carruth et al 1994) the feasibility of this system is not clear. One final point is that Tat may interact with the AP-1 complex (via Jun) to target its transactivation domain to the MVV LTR. It is still unknown whether this Tat/AP-1 interaction could modify the recognition activity of the AP-1 complex.

The experiments in this chapter concentrated on the binding of AP-1 to the EV-1 LTR. The TATA box proximal AP-1 site in the MVV LTRs is in close proximity to a degenerate AP-4 site (Fig.3.5). AP-4/AP-1 interactions at promoters have been described in a number of cellular and viral systems (Mermod *et al.*, 1988, Comb *et al.*, 1988). These two factors act in a synergistic manner when they bind in proximity to each other in promoter and enhancer sequences. The role of AP-4 in the control of the MVV LTR, if it has one, is unknown. DNase footprinting on the 1514 LTR only shows protection over the AP-1 site in this region. Neither the AP-4 nor the VSS sequence are protected (Gabuzda *et al.*, 1989). Further, using a synthetic oligonucleotide covering the AP-4/AP-1 region of the 1514 LTR Neuveut *et al.* (1993) failed to observe an AP-4 specific shift. This suggests that 1514 AP-4 site is non-functional or of low affinity and the requirement for both the AP-4 and AP-1 sequences in Tat transactivation is independent of AP-4 binding. There is the possibility of a non-AP-4, Jun containing, factor recognising this region (Neuveut *et al.* 1993). One problem with the study by Neuveut *et al.* is the short oligonucleotides used as specific competitors. For the AP-4

experiment a 9bp sequence was used and for AP-1 a 12bp sequence. Given the potential problems of using short oligonucleotides in gel shift assays (Section 4.2) and the failure of these workers to show that their 'AP-4' and 'AP-1' sequences did indeed bind these factors this work should be taken as inconclusive. In the EV-1 and SA-OMVV viruses, which have lost their consensus AP-1 sites, the AP-4 site has moved closer to consensus when compared to the 1514 virus (Fig.3.6). The ability of these two viruses to bind AP-4 has not been tested but this may represent a second distinction between the 1514 type viruses (1514 and 1772) and EV-1 and SA-OMVV. While AP-4 and AP-1 can act in a synergistic manner the binding of these two factors to closely positioned sites is independent (Mermod *et al.*, 1988). Thus changes in the recognition sequences should be considered independently rather than as a single unit where the binding of one protein enhances the binding of the second. It appears possible that the interaction of the transcription factor AP-4 may represents another divergence between the 1514 and EV-1 viruses in relation to the mechanisms of transcriptional control. AP-4/AP-1 interactions have also been proposed to regulated FIV transcription (Thompson *et al.*, 1994, Miyazawa *et al.*, 1993). Deletion of this region reduces LTR activity in transfected cells but does not affect the growth of virus *in vitro*. The ability of these deleted viruses to grow *in vivo* has not been reported so it is still possible that this deletion blocks the infective process *in vivo*. This data from MVV together with the observed variation in activity of the MVV LTR variants described in Chapter 3 supports the idea of multiple regulatory elements capable of partially compensating for the loss of any individual sequence. It is also possible that viral transactivators play a role in maximising transcription when the promoter sequence is damaged (Thompson *et al.*, 1994). The possible significance of the variation in transcription factor binding for LTR function will be discussed in Section 5.6.

The gel retardation experiments described in this Chapter provide strong evidence against the interaction of AP-1 with the EV-1 LTRs. The identification of a second, apparently constitutive, DNA binding protein which interacts with the 1514 and EV-1 LTRs sheds more light on regulation of MVV replication. In the course of infection most infected cells will not be in an activated state, and so will be depleted for transcription factors such as AP-1. Lentiviruses present in these cells are believed to be in a state of 'restricted replication' where only the products of spliced transcripts are present and little or no infectious virus is produced (Brahic *et al.*, 1981). The identification of this apparently constitutive factor suggests a mechanism for viral transcription in resting cells.

Gel retardation assays give no information on the function of a DNA binding protein in the context of an active promoter. In order to address this question the recognition sequence for a given factor must be inserted in a vector containing a basal promoter and any *cis* - activating ability assesed. This is important information as different members of the same transcription factor family can regulate transcription in different ways (Majello *et al.*, 1994, Chiu *et al.*, 1989). In order to test the sequences used in these gel retardation assays vectors were generated to examine there ability to upregulate transcription from a basal promoter. These experiments are described in Chapter 5.

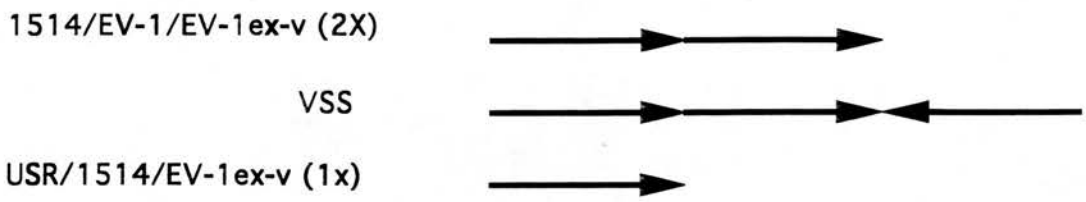
CHAPTER FIVE

5.1 Introduction

The data presented in Chapter 4 demonstrated differences in the binding of transcription factors to the 1514 and EV-1 LTRs. As the identification of binding proteins by the gel shift technique is dependent upon the transcription factors and complexes being stable under assay conditions it is desirable to assess the functionality of these sequences in the context of a basal promoter. This technique demonstrates whether the results observed using the sequences in gel retardation experiments also hold true in the context of a functional promoter. As the sequences under investigation here are believed to interact with transcriptional activators, the assay vector chosen was a minimal promoter, rather than a more extended sequence with a higher basal activity. Such vectors are more suited for testing of 'enhancer' or 'silencer' activities.The vector used for these studies was Δ56. This plasmid contains the CAT gene and the TATA box region of the *c-fos* gene. The presence of a TATA box allows assembly of the basal transcription complex and initiation of transcription at a low rate. Insertion of sequences of interest upstream of this region, using a Sal 1 site, allows them to be tested for any *cis* regulatory function, i.e. do transcriptional activators bind to them. As the basal promoter has a very low background activity any upregulation of transcription is easily detected by reporter gene assays. In order to facilitate cloning into Δ56, the oligonucleotides used in the gel retardation were designed with Sal1 ends, with the exception of the short AP-1 oligonucleotides. The vectors used in the following experiments are outlined in Fig.5.1.

In relation to the factors seen to bind to the MVV LTR there were two points to address. The first was to show that the lack of AP-1 binding to the EV-1 TATA box proximal AP-1 site was a genuine observation and not an artifact of the gel retardation assay. The second was in relation to the second transcription factor seen to interact with the MVV LTR; the VSS binding protein. Is this factor able to function as a transcriptional activator? Comparison of the vectors outlined in Fig.5.1 by transfection into chondrocytes addressess these two questions and tests whether or not the observations made using the gel retardation assay can be extended to transcription processes within the cell. In experiments using the Δ56 vector activity of the oligonucleotide containing vectors is expressed as fold increase over the basic vector (xΔ56).

Figure 5.1 Orientation of oligos in Δ56 Vectors



The arrows denote the orientation of the oligonucleotides relative to the basal promoter. Each arrow represents one copy of the test sequence.

5.2 Comparison of EV-1 and 1514 TATA Box Proximal Regions

The initial experiments focused on whether the differences seen in the gel shift experiments using the long oligonucleotides [1514, EV-1 and EV-1ex-v (Fig.4.1a)] had any bearing on the ability of these sequences to function as *cis* -activators of a heterologous promoter. The differences observed in the gel shift experiment do appear to be reflected in the results of these transfection experiments (Fig.5.2).

These data demonstrate that the 1514 sequence shows the greatest activity whether present in one (Fig.5.2a) or two (Fig5.2b) copies. Where two copies of the 1514 sequence are present an activity 7-8 times that of the EV-1ex-v vectors is seen. These differences are maintained regardless of whether the cells are stimulated by either 5% FCS or PMA (100ng/ml). For the PMA induction experiments cells were cultured in media containing 0.5% FCS. Three hours before harvesting this was replaced by serum free media containing 100ng/ml PMA.

The reduced activity of the EV-1ex-v sequence may be accounted for in one of two ways. If the AP-1 site is responsible for the activity seen then either the AP-1 site is of a lower affinity than the 1514 site, and so less active, or interacting with a different, less transcriptionally active AP-1 complex. The second possibilty is that the AP-1 site within the EV-1ex-v sequence is indeed inactive, confirming the results of the gel shift assays (Chapter 4), and the activity of the vector is accounted for by the factor responsible for the non-AP-1 shift in the gel retardation experiments. Comparison of the activity of the EV-1 and EV-1ex-v sequences (Fig.5.2.) gives some support to the idea that the non-AP-1 factor is responsible for the activity of the EV-1 vectors. Although the difference is not as marked as seen in the gel shift experiments, the EV-1 sequence does appear to show a reduced activity in relation to the EV-1ex-v sequence. As these two sequences differ only outside the AP-1 site and within the VSS sequence, it appears likely that this region is responsible for the activity observed.

Where only a single copy of the oligonucleotide is present, the 1514 sequence again shows an elevated activity in comparison to the EV-1ex-v sequence. However, in these single copy vectors, the difference in activity is less marked suggesting that while multiple AP-1 binding elements function synergistically multiple copies of the EV-1ex-v sequence show only an additive effect.

FIGURE 5.2 Comparison of Δ56 vectors

The relative activity of Δ56 vectors is compared. All experiments show the results from 4 independent transfections with the exception of EV-1 (2x) in Fig.5.2c which is for three independent transfections. Graphs show:

a) one copy vectors in 5% FCS stimulated chondrocytes.

b) two copy vectors in 5% FCS stimulated chondrocytes.

c) two copy vectors in PMA (100ng/ml) stimulated chondrocytes.

The activity of vectors (xΔ56) represents the increase in activity of the MVV Δ56 vectors over the activity of the Δ56 vector without any inserted sequence.

Figure 5.2a

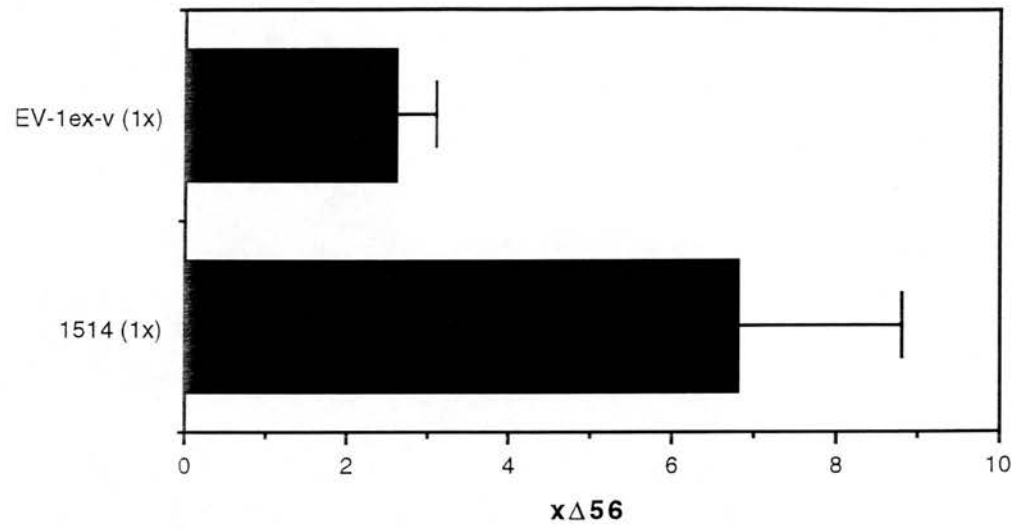


Figure 5.2b

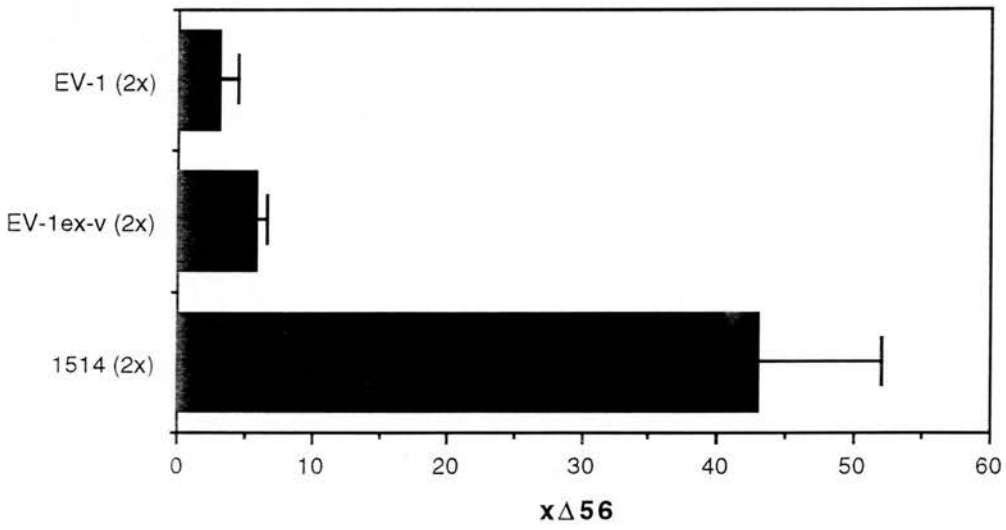
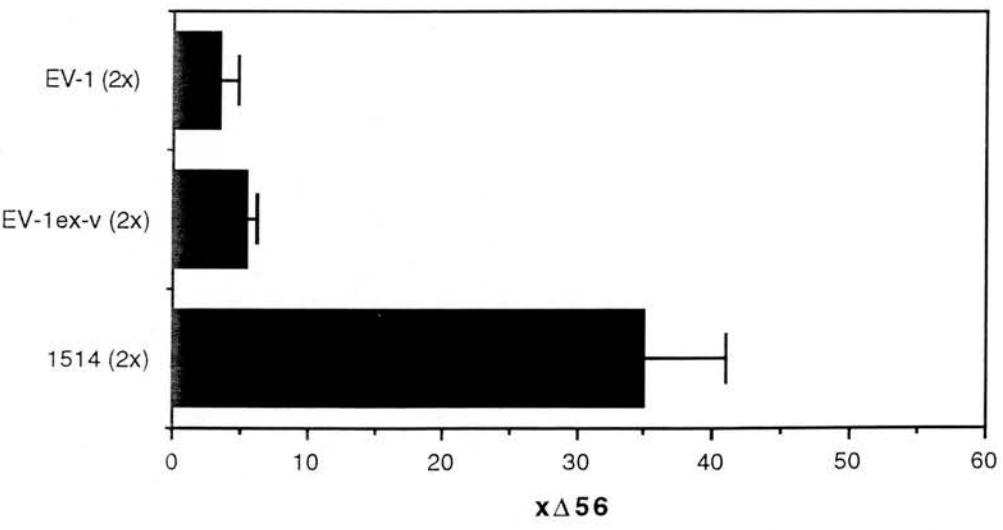


Figure 5.2c



These results are compatible with the observations made in Chapter 4 where each of these 3 sequences showed distinct characteristics in the gel retardation assays. While the experiments described above provide indirect evidence for the VSS sequence binding a transcriptional activator, the question can only be answered using vectors containing the VSS sequence in the absence of an AP-1 site.

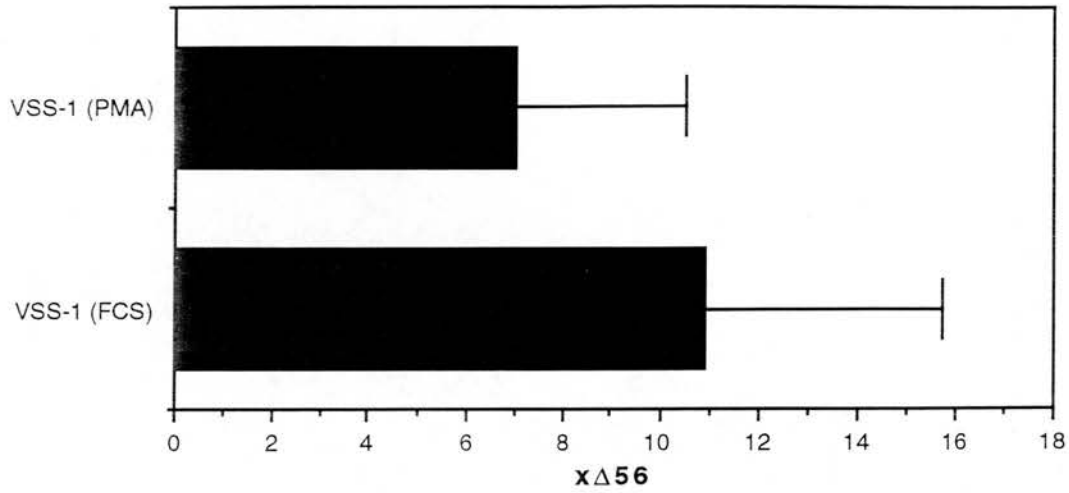
5.3 The Factor Which Binds the VSS is a Transcriptional Activator

As analysis of the long oligonucleotides is complicated by the possible presence of two transcription factor binding sites. The oligonucleotide VSS, which has been shown to only interact with a non-AP-1 factor (Fig.4.5), was used to test whether this factor was indeed a transcriptional activator.

The results of transfections experiments using this vector are shown in Fig.5.3. These data indicate that this sequence is indeed capable of driving transcription from a basal promoter. Again, as for the vectors containing both the AP-1 and VSS elements, the activity is comparable following serum or PMA stimulation. This would support the idea that in the EV-1 viruses the TATA box proximal AP-1 site is non-functional and that transcriptional activation caused by this area of the LTR is due, at least in part, to the binding of a factor to the VSS sequence. The apparent failure of AP-1 to bind to this EV-1 sequence and the lack of a consensus AP-1 site elsewhere within the EV-1 LTRs, raises questions about the studies on the 1514 LTR and Tat function. Studies on the regulation of transcription in 1514 have repeatedly highlighted the importance of AP-1 binding for induction of transcription and transactivation by Tat (Gabuzda *et al.*, 1989, Hess *et al.*, 1989, Shih *et al.*, 1992, Gdovin & Clements 1992, Neuveut *et al.*, 1993).

As the oligonucleotide carrying solely the VSS sequence was capable of driving transcription and the results from the the gel shift experiments suggested a lower affinity for the EV-1 VSS site when compared to the EV-1ex-v and 1514 sequences, the activity of this sequence in the context of the LTR was examined. EV-1 LTRs no.1 and 48 differ in that LTR 48 carries the EV-1ex-v sequence for the VSS binding factor (Fig.3.4).

Figure 5.3

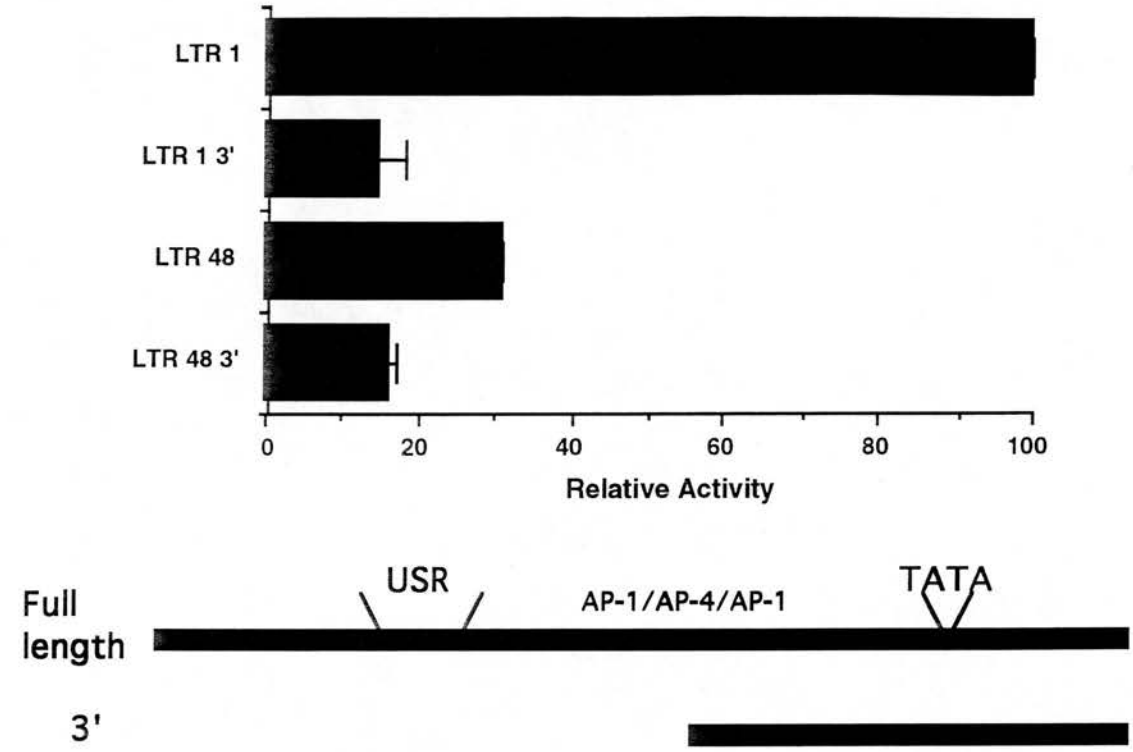


The VSS-1 vector transfected chondrocytes were stimulated with either 5% FCS or PMA (100ng/ml). Activity is the increase over the Δ56 vector without any insert. Data is the mean of 5 (5% FCS) or 4 (PMA) independent transfections.

In order to test the activity of this region, in the absence of upstream regulatory elements, a sequence from position 120 was amplified by PCR and cloned into pCAT 12. Transcription from these two vectors was assessed. The results of these experiments are summarised in Fig.5.4. These two sequences show identical activity demonstrating that in the context of the LTR the G to A conversion within the VSS probably does not signficantly reduce the rate of transcription from the LTR. Comparing the activity of these two vectors with the full length sequences from which they are derived it is clear that the TATA proximal AP-1/VSS element is not causing the reduced activity of LTR 48 in comparison to LTR 1 (Fig.5.4). This would suggest that the difference in activity seen between LTRs 1 and 48 is caused by differences in upstream elements. These two LTRs differ at a number of positions in the upstream sequence due to point mutations (see Fig.3.4). However, the more significant difference between these two sequences is probably the insert at position 90 in LTR 48 (see Fig.3.8). In Section 3.3.2 it was demonstrated that this sequence determined the ability of repeats to act as upregulators or inhibitors of LTR activity. One possible explanation is that this sequence, or a factor which may bind to it interferes with the ability of upstream sequences to contribute to transcriptional activation. It may, in a sense, function as an isolator element preventing sequences upstream contributing fully to promoter activity. Duplications downstream of this site are not effected by the sequence and so allow transactivation (LTRs 18 and 44 compared to LTR 48, Fig.3.5). When this sequence is removed, upstream elements (Section 5.4) can interact more fully with the transcription complex assembling over the TATA box and so upregulate transcription. The presence of repeats, when the position 90 insert is not present, may now interfere with promoter architecture, perhaps by blocking upstream sites, and so lead to a reduction in LTR activity. In this case the repeat is functioning as an analogue of the position 90 insert (LTRs 19 and 30 compared to LTR 1 and also LTR 48). This mode of action is of course highly speculative but does provide one possible explanation for the experimental observations.

In the previous chapter the issue of whether or not both the AP-1 site and VSS could be occupied simultaneously was discussed. While it was not possible to give a definitive answer to this question it appeared unlikely that both sites could be occupied at once. Bringing this observation together with the data on the activities of the various oligonucleotide Δ56 vectors it appears that AP-1 may be capable of displacing the VSS binding protein. This is based on the data in Fig.5.2 where it is seen that the vectors containing the 1514 sequence

Figure 5.4



Activity of truncated LTRs. Vectors containing LTRs 1 and 48 with the upstream sequences deleted were compared by transfection into chondrocytes growing in 5% FCS. Activity is compared to the full length LTR 1 (100). In contrast to the full length LTRs these 3' sequences show comparable activity. The activity of the pCAT12 vector with no LTR sequences was >1. Data is the mean of four independent transfections.

show the highest activities. In the gel shift experiments the VSS binding protein interacted with both the 1514 and EV-1ex-v with an apparently equal affinity. Yet the 1514 sequence is more active in transfection experiments, suggesting that AP-1 binding to 1514 accounts for the higher activity. This AP-1 interaction with the 1514 LTR is probably at the expense of the VSS binding protein. Further experiments are required to answer this question fully; for example, would mutation of the VSS in the 1514 oligonucleotide reduce its activity in Δ56 in comparison to the wildtype sequence? While some questions remain about the interaction of these two factor binding sites in this short sequence it appears that in 1514 at least the ratio of these two factors in the nucleus may play a pivitol role in setting the rate of viral transcription.

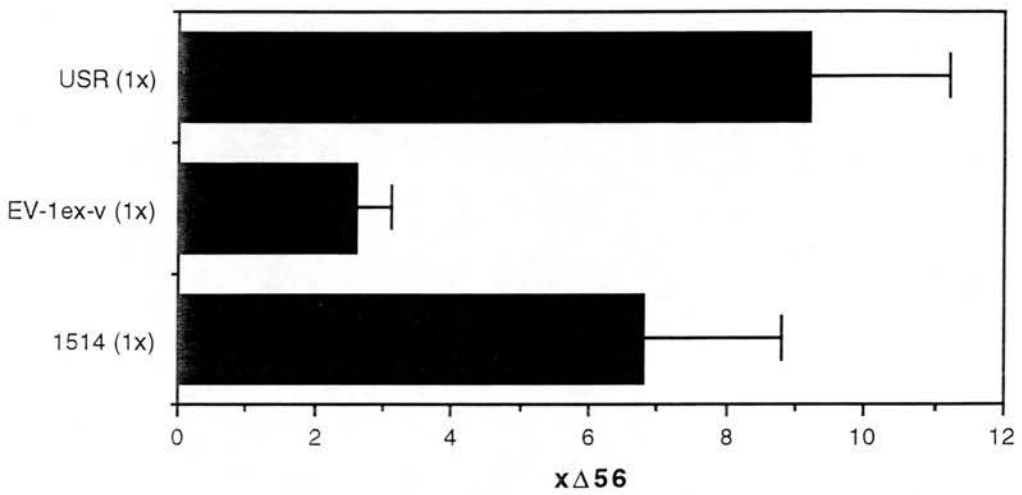
5.4 The USR Sequence From MVV 1514 can also Activate Transcription

The upstream region of MVV 1514 was also inserted into Δ56 as a single copy in the same orientation as present in the viral LTR. This sequence was shown in gel shift experiments to contain a single copy of the VSS. In transfection experiments this sequence is able to activate transcription but shows a distinct activity in comparison to other regions containing only a single copy of the VSS (Fig.5.5). The activity of the USR region is elevated when compared to a single copy of the EV-1ex-v sequence and is equal to that shown by the 1514 sequence which contains a functional AP-1 site. The single copy of the USR sequence appears to drive transcription as well as three copies of the VSS oligonucleotide.

This difference in activity between the USR and EV-1ex-v sequences, which in gel shift analysis appeared to bind the same factor, may be explained in two ways. Firstly, the USR sequence may contain sites for factors not observed in the gel shift analysis, the consensus E-box site being a candidate. Alternatively, the context of the VSS core sequences in these two regions results in the USR site having a greater activity. In terms of context the sequences flanking the VSS are likely to be critical. These do differ between the two VSS sites in 1514 LTR (Fig.4.9). It is of course possible for both of these mechanisms to be functioning.

The data from experiments performed on the VSS and USR sequences argues that the factor observed to bind this region (Chapter 4) is indeed a transcriptional activator. In Chapter 4 the relationship between the VSS in the MVV LTRs and the box 4 footprint in the TCR alpha gene enhancer (Tα4) was discussed. The Tα4 sequence has been shown to footprint but does not appear able to act as an enhancer element (Ho *et al.*, 1989, Ho &

Figure 5.5



Activity of the USR sequence in the vector Δ56.

Transfections were performed on chondrocytes growing in 5% FCS. The USR activity is the mean of 8 independent transfections. The EV-1ex-v and 1514 data is the mean of 4 independent transfections.

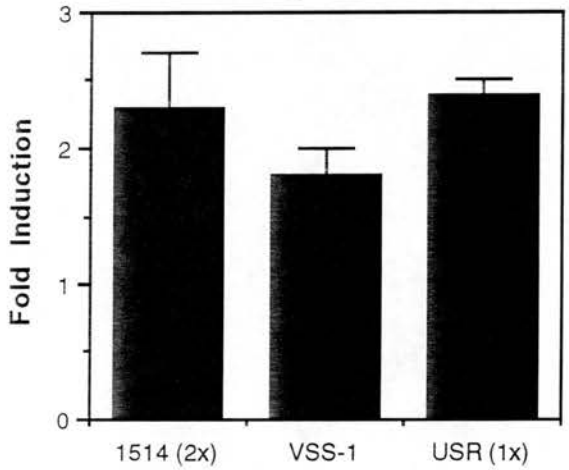
Leiden 1991). The data presented here show that a factor which interacts with this sequence can mediate the activation of transcription. These two observations are not contradictory. As enhancers are positioned at remote points from the promoter they must act at a distance (Chapter 1.7). This action will to some degree depend on the factors ability to remodel chromatin structure. The factor(s) may be acting either indirectly to alter conformation and structure of the chromatin and so allow other factors access to the promoter or directly by bringing enhancer elements into contact with the promoter, via DNA bending or looping. There is experimental evidence to support the idea that not all transcription factors are capable of functioning as enhancers and that this effect is in part mediated by the activation domain (Seipel *et al.*, 1992).

5.5 Serum Induction of Δ56 Vectors

In Section 3.6 serum induction of full length LTRs was described. These experiments showed a uniform (2 fold) induction. The results on induction of AP-1 and the VSS binding factor (Section 4.7) showed that unlike AP-1 the factor binding to the VSS did not appear to be inducible. As there will be multiple transcription factor binding sites within the LTR it was decided to repeat these serum induction experiments using the Δ56 vectors. These experiments were performed using the vectors 1514 (2x), VSS-1 and USR (1x) and the protocol described in Section 3.6. The results from these induction experiments are shown in Fig.5.6. As in the previous attempt (Chapter 3) to show serum induction all three vectors show an equivalent response.

In the gel retardation experiments (Chapter 4) AP-1 activity was present at easily detectable levels in serum starved cells. Thus it is possible that in the system used, AP-1 concentrations do not become limiting and there is a sufficient pool of the factor to activate AP-1 sites within promoters, such as the 1514 LTR or sequences derived from it. It therefore seems likely that the induction observed is not due to specific transcription factors but a general increase in transcription triggered by the elevated serum. In this case factors, such as AP-1 and the VSS binding protein, are already saturating prior to the serum induction.

Figure 5.6



Serum Induction of Δ56 vectors.

The increase in activity of vectors in cells growing in 5% FCS versus 0.5%. In each case the induction was performed in three independent transfections.

5.6 Discussion

The data presented here strongly argues that studies on the regulation of transcription in MVV strain 1514 cannot be extrapolated to describe the the regulation of transcription in MVV strain EV-1 (and by extension possibly SA-OMVV).

The data presented in this Section taken together with the previous studies on AP-1 binding to the MVV LTR (Chapter 4) and the degenerate nature of the AP-1 sites in the EV-1 LTRs, would argue against functional AP-1 sites in the EV-1 LTRs derived *ex vitro* or *ex vivo*. As AP-1 has been proposed as the main regulator of inducible transcription from the MVV LTR and the principle target for Tat transactivation, this result was surprising. As all the remaining AP-1 sites within the EV-1 site are highly degenerate with only limited homology to the consensus sequence, it appears likely that transcription from the EV-1 LTRs is regulated by a system distinct from that used by the 1514 LTRs. The identification of a second factor binding in this region of the LTR in both EV-1 and 1514 viruses suggested a possible mechanism for the replacement of AP-1. However, from the data presented here it appears that this second factor is not related to AP-1 and shows a weaker transactivation function. One problem with the system used here was the failure to show induction of transfected AP-1 sites by serum (Fig.5.6). From the data presented in Chapter 3 this may reflect the 'activated' nature of chondrocytes in this culture system. The data on the truncated LTRs (Fig.5.4) are in approximate agreement with the observations of Hess *et al.* (1989) at least for the LTR 1 sequence, in that deletion of the upstream sequences results in a sharp drop in LTR activity. In contrast LTR 48 appears to suggest that the upstream elements are redundant. This is unlikely to be true and probably reflects the experimental system used. One experiment which remains to be performed is a comparison of the truncated 1514 sequence with the LTR 1 and LTR 48 sequences. This would show the effect of the consensus AP-1 site in the context of LTR sequences.

If these observations are valid in other systems and the VSS binding factor is constitutive then this raises two distinct questions; how do the EV-1 LTRs regulate inducible expression and what is the function of this second factor if it does not confer inducibility on the MVV LTR? If the observations for the induction experiments in chondrocytes hold true for the more relevant case of cells of the monocyte/macrophage lineage then it is possible that the VSS binding factor may drive low level transcription and play a role in restricted replication.

The VSS binding factor(s) may be capable of driving low level transcription from the LTR so building up a pool of Tat and Rev protein. This may provide a mechanism for gearing the infected cell up for high level replication following cellular activation; both Tat and Rev would be at optimal or near optimal concentrations allowing immediate high level production of transcripts for structural proteins. It is likely that there will be a number of other regulatory factors interacting with the LTR and controling its activity. A fuller understanding of the control mechanisms of the MVV LTR will depend on the identification of such elements.

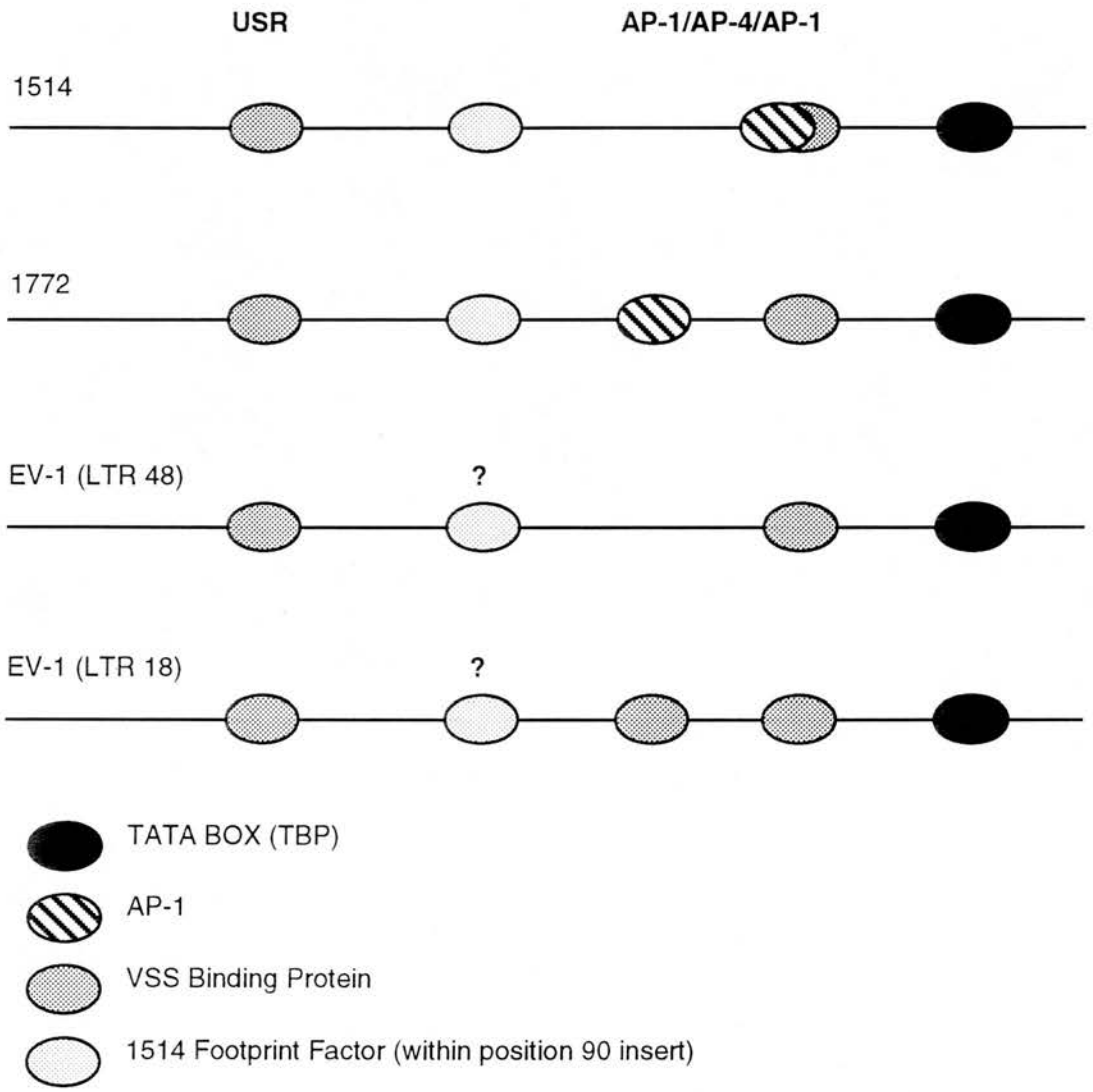
Considering the entire structure of the MVV LTR then it becomes apparent that there is some plasticity in LTR structure. Comparing the MVV LTRs, both variation in transcription factor binding (1514 compared to EV-1) and alterations in the positions of transcription factor binding sites (1514 compared to 1772) are seen. These three LTR types appear likely to possess distinct LTR structures due to variations in transcription factor binding. Possible promoter/transcription factor complexes are outlined in Fig.5.7. It appears a strong possibility that MVV strains are able to alter their transcriptional control mechanisms without greatly altering the *in vivo* infectious process. What does this say regarding the importance and significance of this variation in promoter structure? The main conclusion would appear to be that the exact organisation of the promoter is not under selective pressure. Rather any sequence and architecture which allows the correct rate of transcription for the infective process is tolerated. This lack of constraint on LTR structure is illustrated by the observation that the LTR appears to be as variable as the env gene and more variable than the gag gene (R. Zanoni pers. communication). Both Env and Gag are under strong selective pressure from the immune system so variation is being driven rather than drifting, as in the LTR. However, the plasticity of LTR structure probably has defined limits. This appeared to be the case in the types of LTR selected from the infecting population (Chapter 3); not all LTR types appear capable of supporting *in vivo* infection. Secondly, the data on the transcription rates of the various EV-1 LTR variants (Chapter 3) showed that context was important. Similar observations have been made in other systems. Attempts to modify the FIV promoter by the insertion of 4 additional AP-1 sites into a wild type LTR revealed some constraints on promoter structure (Miyazawa *et al.*, 1994b). Insertion of these AP-1 sites elevated transcription when measured by CAT assay. However, when this modified LTR was introduced into an FIV molecular clone the virus had a reduced rate of replication and after several rounds of infection the AP-1 sites were removed by deletion and a virus with wild type characteristics was regenerated. This

FIGURE 5.7 Transcription binding to MVV LTR types

A schematic representation of transcription factor binding to the MVV LTR is shown. Sequence variation will apparently lead to distinct transcription factor binding patterns in the three LTR types shown. For a comparison of the sequence in AP-1/AP-4/AP-1 regions of the viruses shown see Fig.4.6. In the EV-1 viruses the position 90 insert is shown as a transcription factor binding region given a) its homology to the region footprinted in the 1514 LTR and b) the observations on the LTR variants in Chapter 3 which suggests it modulates LTR activity.

As can be seen from this Figure the four LTRs show distinct transcription factor binding properties. In the 1514 LTR it is still unclear whether both AP-1 and the VSS binding protein can bind simultaneously. However in 1772, with the movement of the AP-1 site, the sites are sufficiently distant (Fig.3.6) to suppose there could be no direct interference, or site competition, between the factors in relation to DNA binding. If there is competition between these sites in 1514 then it appears that the VSS binding protein is displaced by AP-1. Duplication of the AP-1/AP-4/AP-1 region in some of the EV-1 LTRs generates two sites for the VSS binding protein. This duplication approximately doubles LTR activity (Fig.3.5.). The significance of this duplication for virus replication *in vivo* is unclear as both repeatless and repeat bearing LTRs are tolerated in EV-1 infected animals.

FIGURE 5.7


USR
AP-1/AP-4/AP-1
1514
1772
EV-1 (LTR 48)
?
EV-1 (LTR 18)
?
TATA BOX (TBP)
AP-1
VSS Binding Protein
1514 Footprint Factor (within position 90 insert)

would suggest that the LTR has already evolved to its optimum form and attempts to enhance artificially its activity may disrupt the 'stereospecific ' complex which normally forms at the promoter (Chapter 1.7). Together with the observations presented here on the variation in MVV LTR structure it would appear that promoter plasticity only occurs within fixed limits. The boundaries being set by the nature of the interacting factors in a given promoter and there dependency on a given structure.

One further sequence which should be compared between the various MVV LTRs is the TATA box itself (Fig.5.8). In HIV-1 TATA box mutants have been identified as second site revertants which compensate for the loss of upstream Sp1 sites (Kashanchi *et al.*, 1994). Variation in TATA box sequences may also modify the Tat responsiveness of the HIV-1 LTR (Berkhout & Jeang 1992). The TATA box sequences of the various MVV isolates from Fig.5.7 are compared in Fig.5.8. From this comparison it is clear that these sequences are varying between MVV isolates. The significance of this variation is unknown but the point should be made that there may well a relationship between alterations in transcription factor sites in the promoter/enhancer proper and the 'basal' component represented by the TATA box sequence.

The interactions between transcriptional control systems can result in a potentiation or inhibition of transcription. In the 1514 virus the overlapping AP-1 site/VSS represents a target for two distinct sets of factors. Composite elements have been observed in a number of promoters (Diamond *et al.*, 1990, Rahuel *et al.*, 1992, Tan *et al.*, 1992, Treisman 1992, Grazia *et al.*, 1994). The serum response element of the *c-fos* gene represents an extreme example where 7 distinct factors, including the serum response factor, may interact with a 22bp region of DNA (Treisman *et al.*, 1992). Factors interacting with a composite element my compete for binding resulting in a suppression or activation of transcription (Rahuel *et al.*, 1992, Tan *et al.*, 1992, Grazia *et al.*, 1994). Alternatively factors may interact co-operatively to upregulate their binding and so synergisticly activate transcription. In the case of the Interleukin 2 enhancer both these effects are seen (Grazia *et al.*, 1994). Co-operartive binding of AP-1 and Oct-2 results in the activation of transcription. However, the retenoic acid receptor (RRR) binds to the same region of the enhancer and this interaction disrupts the AP-1/OCT complex and inhibits transcription. The effect of factor binding may also depend on the composition of the factor interacting with the site. As has been stated before the factor AP-1 factor may be composed of numerous combinations of the Fos and Jun proteins depending on cell lineage and activation

Figure 5.8 Comparison of TATA box sequences from MVV isolates

1514	GCCTATA**T**AAGCCG
SA-OMVV	GCCTATA**A**AAGCTG
EV-1 *ex vivo*	GCCTATA**A**AAGCTG
EV-1 ex vitro	GCCTATA**A**AAGCTG

The TATA box sequences of 1514 type viruses (Sonigo *et al.*, 1985, Staskus *et al.*, 1991) and the viruses SA-OMVV (Querat *et al.*, 1990) and EV-1 (Chapter 3) are compared. The notable feature of these sequences is the T to A conversion (in bold type) in the viruses which appear to lack TATA box proximal AP-1 sites (EV-1 and SA-OMVV).

state. While this will effect the transcriptional activation potential of the AP-1 protein it will also determine the interaction of AP-1 with other transcription factors (Miner & Yamamoto 1992). Comparison of the plfG element from the proliferin gene promoter binds both AP-1 and the GR. The compostion of the AP-1 complex effects the activity of this sequence; while Fos containing AP-1 complexes repress, complexes containing Fra-1 activate transcription from the element. This effect was not mediated by the activation domain but the DNA binding domain. Thus this interaction is similar to that observed in the HIV LTR between Nf-kB and Sp1 (Perkins *et al.*, 1994). This crosstalk between the AP-1 and hormone receptors has been examined by other groups (Yang-Yeng *et al.*, 1990, Doucas *et al.*, 1991, Grazia *et al.*, 1994). This suggests that these two sets of factors are involved in a complex crosstalk system which is modulated by the set of *fos* and *jun* family members expressed in the responding cell.

Cross talk between different signalling pathways may thus occur at composite elements. Such sequences may represent crucial elements in the control of gene expression. In the 1514 virus there was no evidence *in vitro* for simultaneous binding of AP-1 and the VSS binding protein to the closely positioned TATA box proximal sites. From the transfection data it appeared that the VSS binding factor was being displaced by AP-1. It also appears that the VSS binding factor is unable to recruit AP-1 to the non-consensus AP-1 site present in the EV-1 oligonucleotides. This data strongly suggest that in this region of the LTR there is competition for DNA binding between these two factors. In this regard a comparison of the AP-1/AP-4/AP-1 regions from the 1514 and 1772 viruses would be interesting as in the 1772 LTR the AP-1 site has been repositioned due to point mutations in the LTR. This results in a separation of the AP-1 site and VSS sequence.

CHAPTER SIX

CONCLUSIONS

6.1 Introduction

In vivo the activation of transcription from the MVV LTR appears constrained by cell type and differentiation state (Chapter 1). Within cells of the monocyte lineage differentiation into macrophages, and macrophage activation, is required for the efficient expression of transcripts driven by the MVV LTR. For MVV strain 1514 the transcription factor AP-1 has been proposed as a major regulator of LTR activation. In cells of the monocyte/macrophage lineage differentiation and activation leads to the expression of the *fos* and *jun* gene families (Matsui *et al.*, 1990, Redner *et al.*, 1992). These factors are positive regulators of monocyte differentiation (Lord *et al.*, 1993). Studies on monocytes, and other cell lineages, illustrate the differential regulation of the various *fos* and *jun* family members. Thus, during the timecourse of activation the composition of the AP-1 dimers present within the cell will change. This alteration of dimer composition will have effects on the activation of genes, due to changes in site affinity and activation potential, and also modulate the crosstalk between AP-1 and other transcription factor families.

6.2 Crosstalk and Transcriptional Regulation

Crosstalk between different transcription factor families can obviously occur between independent sites in the target promoter and also by competition or cooperation at composite elements (Section 5.6). In addition to these two mechanisms there is also the possibility of direct protein-protein interactions between members of different families and the competition of different transcripion factor complexes for the same DNA sequence, as opposed to interference at composite elements where two distinct recognition sites are involved.

Competition for a single site is observed between a number of factors as well as within families (i.e. the varying site affinities and activities of different AP-1 forms). Cross family competition is observed at AP-1 and CRE (cAMP response element) sites. The recognition sites of these factors are closely related and allows these two factors to recognise reciprocally each others consensus sites, but with distinct affinities (Galien *et al.*, 1994, Ryseck & Bravo 1991). Further, the affinity of AP-1 for the CRE is dependent on the composition of the AP-1 dimer. These two factors may show distinct effects on promoters containing an AP-1 site or CRE; CREB potently inhibits the Ras driven activation of the collagenase TRE (TPA response

element, a consensus AP-1 site) (Galien *et al.*, 1994). Thus an AP-1 site (or CRE) may be targeted by members of both the Fos and Jun families (AP-1) and CREB (a member of the ATF-CREB familiy) and the consequences in terms of transcriptional regulation will be dependent on which factor binds the site. The consensus AP-1 site, and CRE, may also be targeted by NF-IL6 (C/EBPß) a member of the C/EBP family (Klampfer *et al.*, 1994, Tsukada *et al.*, 1994). At a single CRE site within the prointerleukin-1 gene 3 distinct factors are seen to bind; NF-IL6 dimers, CREB dimers and NF-IL6/CREB heterodimers. NF-IL6 is present in a preexisting inactive form and is activated by stimuli such as LPS and IL-6. Activation triggers translocation to the nuclei and activation of target genes, including *IL-6* and *c-fos* (Akira *et al.*, 1990, Metz & Ziff 1991). NF-IL6 is composed of positive and inhibitory isoforms which are translated from the same RNA species from separate in frame AUGs (Descombes & Schibler 1991, Hsu *et al.*, 1994). The inhibitory isoform is capable of binding to some AP-1 sites so blocking AP-1 binding and inhibiting gene transcription (Klampfer *et al.*, 1994). This recognition is not fully reciprocal as AP-1 (Fos/Jun) will not bind most NF-IL6 sites. In addition to this promiscuous site recognition there is the phenomenon of cross family dimerisation.

6.2.1 Interactions between members of different transcription factor families

The example of cross dimerisation between CREB and NF-IL6 has already been mentioned. This is not a unique example. Members of the three transcription factor families Fos/Jun, ATF-CREB and C/EBP all contain a leucine zipper domain which mediates the dimerisation of these factors (Hai & Curran 1991, Lamb & McKnight 1991, Johnson & McKnight 1989). Members of these families are able to cross dimerise selectively with members of the other groups. The basis for this selective cross dimerisation is unclear but the composite factors show distinct site recognition (Hai & Curran 1991, Hsu *et al.*, 1994). Cross dimerisation is limited between families. For example, while ATF-4 will dimerise with Jun, Fos and Fra-1, ATF-1 will not dimerise with any of these proteins and ATF-2 and ATF-3 will only dimerise with Jun (Hai & Curran 1991). These heterodimers show distinct binding specificity depending on their compostion. While an ATF-3/Jun complex will recognise both the CRE and AP-1 sites an ATF-4/Jun complex binds the CRE but not AP-1 sites.

The examples discussed above are from what may be described as a transcription

factor super family (Hai & Curran 1991); all these factors dimerise via a leucine zipper domain. The formation of cross family heterodimers is not dependent on this sharing of a common dimerisation domain as is demonstrated by the cross dimerisation between the Fos/Jun proteins and NF-κB p65 (Stein *et al.*, 1993). This dimerisation is mediated by an interaction between the bZIP domain of Fos and Jun and the Rel homology domain of p65. The interaction of Fos and Jun with NF-κB appears to enhance the binding of these factors to their recognition site *in vitro*. Indirect evidence for these dimers existing *in vivo* comes from data where expression of antisense *jun* or *fos* inhibits the transcription from an artificial promoter driven by mutiple NF-κB sites (Stein *et al.*, 1993). Overexpression of Jun or Fos (but not Jun B or Jun D) potentiated transcription from the same artificial promoter or a truncated HIV-1 LTR (deleted for sequences upstream of the NF-κB sites). Such transfection studies cannot rule out indirect effects but together with the *in vitro* data support the idea of the formation heterodimeric species. Using immunoprecipitation NF-kB p50 has been shown to be able to interact with NF-IL6 apparently also via the Rel and bZIP domains (LeClair *et al.*, 1992). In this case there was no examination of the functional consequences of this association.

This cross dimerisation allows changes in both sequence specificity and the transactivation domains targeted to a given promoter. It provides a mechanism whereby distinct signal transduction pathways may interact. The common occurance of cross dimerisation between transcription families, generating novel activities, and competition of various factors for the binding to the same region of DNA results in a huge number of possible factor combinations at a promoter. During cellular activation there will be a flux in the concentration and types of transcription factors available to interact and regulate transcription. It is therefore likely that over the course of the activation process the composition of the promoter complex will alter as factors are displaced by others with higher affinity for that region of the promoter. In this respect the transcription of the MVV LTR induced by macrophage activation is likely to be dependent on multiple interacting factors. These factors may be interacting at the promoter itself or prior to promoter binding. It will be of interest to determine whether their is direct protein-protein interaction between the factors interacting with the MVV LTR. The work of Neuveut *et al.* (1993) suggests that in 1514 the c-Jun protein may be a component of a non-AP-1 factor interacting with the LTR.

6.3 Summary and Concluding Remarks

The data presented in this thesis argues that different MVV isolates show modified mechanisms of transcriptional regulation. This conclusion is based on the observation that the EV-1 LTRs lack a consensus AP-1 site and that the positions of transcription factor sites are not fixed between the EV-1, 1514 or 1772 viruses. Both 1514, and its derivative 1772, contain a consensus AP-1 site and in the case of 1514 this sequence has been demonstrated to bind AP-1. The sequence within the EV-1 LTRs with closest homology to the consensus AP-1 sequence did not bind AP-1 *in vitro* and did not appear to do so *in vivo*. It was not possible to conclude for certain that this sequence was non-functional *in vivo* due to the presence of a second transcription factor site in this region of the LTR. This factor, which binds the same sequence in the 1514 and EV-1 viruses, shows a transactivating activity. This activity was lower than that seen for the AP-1 complex present within these cells. There appear to be two sites for this factor within the MVV LTR, one proximal to the TATA box the other at an upstream site. This factor appears to be distinct from AP-1 both on the basis of site recognition and the failure to observe binding of an anti-Jun antibody in gel supershift experiments. UV crosslinking experiments identified a protein of approximately 34-40kD which interacts with the recognition site. A full characterisation of this factor, and other cellular factors with which it interacts, will allow further dissection of the mechanism by which MVV controls its transcription rate.

Various questions need to be addressed regarding the nature of the VSS binding protein. The simplest involves determining the tissue distribution of this factor; is there any evidence for tissue specific expression? A second related question is whether this factor is a component of a multi member family; do the family members show tissue specific expression? In relation to MVV the key question is the expression of this factor during monocyte/macrophage differentiation and activation. The transcription factor AP-1 is activated during this process and it is important to determine if the factor which binds the VSS is under similar control.

An analysis of the expression of transcription factors involved in the regulation of MVV during monocyte/macrophage differentiation may also explain the restriction/latency of MVV in the majority of infected cells *in vivo*. *In vitro* the MVV LTR has been shown to be a strong promoter which is active in a variety of different cell types. This is in marked contrast to

the observations of *in vivo* infection where viral transcription appears under tight control. The basis of this *in vivo* restriction is unclear. It is possible that the VSS binding protein is involved in this process, however, nothing is known about its expression *in vivo*. In addition to these two factors, AP-1 and the VSS binding protein, the nature of other factors interacting with the LTR remains to be defined. It is possible that neither of these two factors is critical for the regulation of MVV transcription and some additional factor is required.

One important question raised in this thesis which needs to be fully resolved is the apparent variation between the EV-1 and 1514 viruses in relation to AP-1 binding. Sequence comparison suggested that none of the non-consensus AP-1 sites within the EV-1 LTR would be able to bind AP-1 but this still requires formal proof. The EV-1 TATA box proximal AP-1 site was shown to be unable to bind AP-1 *in vitro*. From the data using the gel shift oligonucleotides in a CAT reporter vector it was not possible to completely rule out AP-1 binding to the EV-1 sequence *in vivo* due to the presence of the VSS binding site. However, this data does indicate that even if AP-1 is binding to the EV-1 sequence it is binding with a much lower affinity than in 1514.

In relation to other factors which interact with the LTR the nature of the factor interacting with the sequence at position 90 remains to be defined. From the data on the EV-1 LTR variants this factor appears to be playing a role in determining the activity of duplications of *cis*-regulatory sequences in the LTR. The effect of this sequence illustrates the codependence of the factors driving transcription from the LTR. As has been discussed crosstalk can involve direct protein-protein interactions or occur indirectly through effects on DNA conformation and the factors which are recruited to the basal transcription complex assembled on the TATA box. One factor which may be affecting LTR activity is integration state. To date studies on the MVV LTR have all made use of transient transfection assays where the DNA is present in a non-integrated form. In contrast during infection the viral genome will be integrated into the hosts DNA. It is possible that this integrated state may affect viral transcription and play a role in latency. Viral latency *in vivo* is likely to be the product of several interacting regulatory pathways. A comparison of the factors interacting with the LTR during latency and active transcription will be required.

In Chapter 3 it was seen that alterations in the structure, and sequence, of the MVV LTR modified its transcription rate. This data illustrated the interaction between various sequences within the LTR and the variability of the LTR sequence within the EV-1 virus

population. Comparisons of the LTRs from the 3 viruses 1514, 1772 and EV-1 suggests that alterations in LTR architecture are tolerated. These three viruses appear to have transcription factors binding in distinct locations in the LTR. A question which remains to be answered is whether such alterations in LTR structure alter the process of viral pathogenesis *in vivo*. The 3 strains do differ in the pathology they induce; EV-1 the British isolate causes predominatly Maedi disease, in contrast the 1514 and 1772 viruses cause primarily Visna disease, with the 1772 virus being selected *in vivo* by serial passage for increased neuropathogenesis. Of course, these viruses do not solely differ in their LTR sequences so the importance of this variation to the distinct pathologies induced remains unclear. Comparing the 1514 and 1772 LTRs then it could be suggested that the separation of the VSS and the AP-1 site may result in an elevated transcription, and replication rate, so accelerating the disease course. This separation of these two binding sites also raises the question of synergistic interaction between the two factors. It is still unclear whether both these factors interact simultaneoulsy at the TATA box proximal site in 1514. This could be addressed using a Δ56 vector containing the AP-1/VSS region but with mutations in the VSS region. The activity of this sequence could then be compared with the 1514 and EV-1ex-v (AP-1 mutant) vectors. Such a comparison should identify any interaction between AP-1and the VSS binding protein which modifies the transcription rate.

One method for determining the role of the LTR in determining the disease course would be by constructing chimeric molecular clones differing only in the LTR sequence. Such constructs could then be introduced directly into animals by DNA injection. The data on the EV-1 LTR population within the infected animal and the comparison of these sequences with those seen in the infecting population strongly suggests that the LTR is under selective pressure *in vivo*. This selection appears to limit the number of LTR types capable of establishing infection. Dissection of the interactions involved in the regulation of transcription from these promoters will depend on targeted disruption of transcription factor binding sites and the manipulation of sites by altering orientation and spacing. While these studies may be performed using isolated LTRs and reporter gene assays, in order to shed light on the LTRs role in disease, it will also be neccessary to perform such experiments using infectious molecular clones. This will allow the effect of alterations of LTR structure on viral replication to be monitored.

One aspect of the regulatory process not touched on in this work is the activity and

targets of the MVV Tat protein. It appears that the MVV Tat protein does possess a potent transcriptional activation domain (Carruth *et al.*, 1994). The absence of a TAR region in the MVV viruses has led to the proposal that the Tat protein is interacting with cellular factors and is activating transcription via this route. A second observation on the MVV Tat protein, that it only weakly transactivates the MVV LTR, has been used to suggest that Tat may be primarily targeting cellular genes so maintaining the cell in an activated state and disrupting normal homeostasis. This mechanism of action is consistent with observations that MVV Tat is not required for *in vitro* replication. It is possible that it is more important for the successful maintenance of infection *in vivo*, and may be involved in the generation of pathology. This remains to be tested by experimental infection with Tat deleted virus.

Due to the apparent requirement for integration in the lentiviral lifecycle these viruses behave essentially as cellular genes. However they have the great advantage of containing all their regulatory elements within a defined sequence of DNA which can be readily manipulated *in vitro*. In contrast, cellular genes contain widely dispersed regulatory elements and these elements cannot be easily manipulated in their normal context. Analysis of the molecular basis of replication in MVV, and other lentiviruses, will shed light on the mechanisms by which these pathogens maintain infection and evade the host immune response. Further, a fuller understanding of the mechanisms involved in the control of lentiviral transcription is useful in the dissection of cellular transcriptional control mechanisms.

BIBLIOGRAPHY

Abate C., Patel L., Rauscher III, F.J. & Curran T. (1990) Redox regulation of Fos and Jun DNA-binding activity in vitro. *Science* **2 4 9**:1157-1161

Abate C., Baker S.J., Lees-Miller S.P., *et al.* (1993) Dimerisation and DNA binding alter phosphorylation of Fos and Jun. *Proc. Natl. Acad. Sci. USA* **9 0**:6766-6770

Adya N., Zhao L.J., Huang W. *et al.* (1994) Expansion of CREBs DNA recognition specificity by Tax results from interaction with ala-ala-arg at position 282-284 near the conserved DNA binding domain of CREB. *Proc. Natl. Acad. Sci. USA* **9 1**:5642-5646

Agamemmnon E.G., Wang Z-Q., Cecchini M.G. et al. (1994) c-Fos: A key regulator of osteoclast-macrophage lineage determination and bone remodeling. *Science* **2 6 6**:443-448

Akira S., Isshiki H., Sugita T. *et al.* (1990) A nuclear factor for Il-6 expression (NF-Il6) is a member of the C/EBP family. *EMBO J.* **9**:1897-1906

Andresson O.S., Elser J.E., Tobin G.J. *et al.*(1992) Nucleotide sequence and biological properties of a pathogenic proviral molecular clone of neurovirulent visna virus. *Virology* **1 9 3**:89-105

Antoni B.A. Stein S.B. & Rabson A.B. (1994) Regulation of human immunodeficiency virus infection: implications for pathogenesis. *Adv. Vir. Res.* **4 3**:53-145

Archer T.K., Lefebvre P., Wolford R.G. & Hager G.L. (1992) Transcription factor loading on the MMTV promoter: a bimodal mechanism for promoter activation. *Science* **2 5 5**:1573-1576

Audoly G., Sauze N., Harkiss G. *et al.*(1992) Identification and subcellular locatization of the Q gene product of visna virus. *Virology* **1 8 9**:734-739

Auwerx J. & Sassone-Corsi P. (1991) IP-1: a dominant inhibitor of Fos/Jun whose activity is modulated by phosphorylation. *Cell* **6 4**:983-993

Baiyong L., Adams C.C. & Workman J.L. (1994) Nucleosome binding by the constitutive transcription factor Sp1. *J.Biol. Chem.* **2 6 9**:7756-7763

Bannister A.J., Gottlieb T.M., Kouzarides T. & Jackson S.P. (1993) c-Jun is phosphorylated by the DNA-dependent protein kinase in vitro; definition of the minimal kinase recognition motif. *Nucleic Acids Res.* **2 1**:1289-1295

Bannister A.J., Brown H.J., Sutherland J.A. & Kouzarides T. (1994) Phosphorylation of the c-Fos and c-Jun HOB1 motif stimulates its activation capacity. *Nuc. Acids Res.* **2 2**:5173-5176

Benezra R., Cantor C.R. & Axel R. (1986) Nucleosomes are phased along the mouse ß-major globin gene in erythroid and nonerythroid cells. *Cell* **4 4**:697-704

Berger S.L., Cress W.D., Cress A.C. *et al* (1990) Selective inhibition of activated but not basal transcription by the acidic activation domain of VP16: evidence for transcriptional adaptors. *Cell* **6 1**:1199-1208

Berkhout B., Gatignol A., Rabson A.B. & Jeang K-T. (1990) TAR independent activation of the HIV-1 LTR: evidence that Tat requires specific regions of the promoter. *Cell* **6 2**:757-767

Berkhout B. & Jeang K-T. (1992) Functional roles for the TATA promoter and enhancers in basal and Tat-induced expression of the human immunodeficiency virus type 1 long terminal repeat. *J. Virol.* **6 6**:139-149

Biggin M.D., Bichel S., Benso M., Pirotta V. & Tijan R. (1988) *Zeste* encodes a sequence specific transcription factor that activates the Ultrabithorax promoter in vitro. *Cell* **5 3**:713-722

Binetruy B., Smeal T. & Karin M. (1991) Ha-Ras augments c-Jun activity and stimulates phosphorylation of its activation domain. *Narure* **3 5 1**:122-126

Birnboim H.C. & Doly J. (1979) A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res.* **7**:1513-1518

Blacklaws B.A., Bird P. & McConnell. (1994) Pathogenesis and immunity in lentivirus infections of small ruminants. *in* Cell-Mediated Immunity in Ruminants, eds. Goddeeris, B.M.L. & Morrison W.I., CRC Press.

Boral A.L., Okenquist S.A. & Lenz J. (1989) Identification of the SL3-3 virus enhancer core as a T-lymphoma cell specific element. *J. Virol.* **6 3**:76-84

Boyle W.J., Smeal T., Defize L.H.K. *et al.* (1991) Activation of protein kinase C decreases phosphorylation of c-Jun at sites that negatively regulate its DNA-binding activity. *Cell* **6 4**:573-584

Brahic M., Filippi P., Vigne R & Haase A.T. (1977) Visna virus RNA synthesis *J. Virol.* **2 4**:74-81

Brahic M., Stowring L., Ventura P. & Haase A.T. (1981) Gene expression in visna virus infection in sheep. *Nature* **2 9 2**:240-242

Bresnick E.H., Bustin M., Marsuad V. , Richard-Foy H. & Hager G.L. (1992) The transcripionally active MMTV promoter is depleted for histone H1. *Nucleic Acids Res.* **2 0**:273-278

Braun M.J., Clements J.E. & Gonda M.A. (1987) The visna virus genome: evidence for a hypervariable site in the *env* gene and sequence homology among lentivirus envelope proteins. *J. Virol.* **6 1**:4046-4054

Buckle R., Balmer M., Yenidunya A. & Allan J. (1991) The promoter and enhancer of the inactive chicken beta-globin gene contains precisely positioned nucleosomes. *Nucleic Acids Res.* **1 9**:1219-1226

Buiser R.G., Bambara R.A. & Fay P.J. (1993) Pausing by retroviral DNA polymerases promotes strand transfer from internal regions of RNA donor templates to homopolymeric acceptor templates. *Biochem. Biophys. Acta* **1 2 1 6**:20-30

Buratowski S. (1994) The basics of basal transcription by RNA polymerase II. *Cell* **7 7**:1-3

Byers Kraus V., Inostroza J.A., Yeung K. *et al.* (1994) Interaction of the Dr1 inhibitory factor with the TATA binding protein is disrupted by adenovirus E1A. *Proc. Natl. Acad. Sci. USA* **9 1**:6279-6282

Carey M. (1991) Mechanistic advances in eukaryotic gene activation. *Curr. Op. Cell Biol.* **3**:452-460

Carey N. &Dalziel (1994) Sequence variation in the gp135 gene of maedi visna virus strain EV-1. *Virus Genes* **8**:115-123

Carillo S., Pariat M., Steff A.M. *et al.* (1994) Differential sensitivity of Fos and Jun family members to calpains. *Oncogene* **9**:1679-1689

Carlsson P., Waterman M.L. & Jones K.A. (1993) The hLEF/TCF-1a HMG protein contains a context dependent transcriptional activation domain that induces the TCRa enhancer in T cells. *Genes Dev.* **7**:2418-2430

Carruth L.M., Hardwick J.M., Morse B.A. & Clements J.E. (1994) Visna virus Tat protein: a potent transcription factor with both activator and suppressor domains. *J.Virol.* **6 8**:6137-6146

Carvalho M., Kirkland M. & Derse D. (1993) Protein interactions with DNA elements in variant Equine Infectious Anemia Virus enhancers and their impact on transcriptional activity. *J. Virol.* **6 7**:6586-6595

Celander D. & Haseltine W.A. (1984) Tissue-specific transcription preference as a determinant of cell tropism and leukaemogenic potential of murine retroviruses. *Nature* **3 1 2**:159-162

Cheng-Mayer C., Quiroga M., Tung J.W., Dina D. & Levy J.A. (1990) Viral determinants of human immunodeficiency virus type 1 T-cell or macrophage tropism, cytopathogenicity and CD4 antigen modulation. *J. Virol.* **6 4**:4390-4398

Chiu R., Angel P. & Karin M. (1989) Jun-B differs in its biological properties from, and is a negative regulator of, c-Jun. *Cell* **5 9**:979-986

Clements J.E., Zink C., Narayan O. & Gabuzda D.H. (1992) Lentivirus infection of macrophages. pp589-600 in Macrophage-Pathogen interactions Immunology series Vol 60 Eds Zuilling B.S. & Eisenstein T.K. Marcel Dekker Inc NY,Basel, H-K

Clements J.E., Wall R.J., Narayan O. et al (1994) Development of transgenic sheep that express the visna virus envelope gene. *Virology* **2 0 0**:370-380

Clerici M., Stocks N.I., Zajac R.A. *et al.* (1989) Interleukin-2 production used to detect antigenic peptiden recognition by T-helper lymphocytes from asymptomatic HIV-seropositive individuals. *Nature* **3 3 9**:383-385

Coffin J.M. (1979) Structure, replication, and recombination of retrovirus genomes: some unifying hypotheses. *J. Gen. Virol.* **4 2**:1-26

Coffin J.M. (1990) Retroviridea and their replication, p1437-1500 *In* B.N. Fields & D.M. Knipe (ed.), Virology, 2nd ed. Raven Press, New York

Cohen D.R. & Curran T. (1988) fra-1: a serum inducible, cellular immediate early gene that encodes a Fos related antigen. *Mol. Cell. Biol.* **8**:2063-2069

Cohen D.R., Sinclair A.H. & McGovern J.D. (1994) SRY protein enhances transcription of Fos-related antigen 1 promoter constructs. *Proc. Natl. Acad. Sci. USA* **9 1**:4372-4376

Cohen R.S. & Meselson M. (1988) Periodic interactions of heat shock transcriptional elements. *Nature* **3 3 2**:856-858

Colotta F., Polentarutti N., Sironi M. & Mantovani A. (1992) Expression and involvement of c-fos and c-jun protooncogenes in programmed cell death induced by growth factor deprivation in lymphoid cells *J. Biol. Chem.* **2 6 7**:18278-18283

Comb M., Mermod N., Hyman S.E. et al (1988) Proteins bound at adjacent DNA elements act synergistically to regulate human proenkephalin cAMP inducible transcription. *EMBO J.* **7**:3793-3805

Conaway C.C. & Conaway J.W. (1993) General initiation factors for RNA polymerase II. *Annu. Rev. Biochem.* **6 2**:161-190

Corboy J.R., Buzy J.M., Zink C. & Clements J.E. (1992) Expression directed from HIV long terminal repeats in the central nervous system of transgenic mice. *Science* **2 5 8**:1804-1808

Courey A.J., Holtzman D.A., Jackson S.P., Tijan R. (1989) Synergistic activation by the glutamine rich domains of human transcription factor Sp-1. *Cell* **5 9**:827-836

Craigie R. (1992) Hotspots and warm spots: integration specificity of retro elements. *T.I.G.* **8**:187-190

Crane S.E., Buzy J. & Clements J.E. (1991) Identification of cell membrane proteins that bind visna virus. *J. Virol.* **6 5**:6137-6143

Crawford T.B., Adams D.S., Cheevers W.P. & Cork L.C. (1980) Chronic arthritis in goats caused by a retrovirus. *Science* **2 0 7**:997-999

Cullen B. R., Skalka A.M. & Ju G. (1983) Endogenous avian retroviruses contain deficient promoter and leader sequences. *Proc. Natl. Acad. Sci. USA* **8 0**:2946-2950

Dalziel R.G., Hopkins J., Watt N.J. et al. (1991) Identification of a putative cellular receptor for the lentivirus visna virus. *J. Gen. Virol.* **7 2**:1905-1911

Davis J.L., Molineaux S. & Clements (1987) Visna virus exhibits a complex transcriptional pattern: one aspect of gene expression shared with the aquired immunodeficiency syndrome retrovirus. *J. Virol.* **6 1**:1325-1331

Davis J.L. & Clements J.E. (1989) Characterisation of a cDNA clone encoding the visna virus transactivating protein. *Proc. Natl. Acad. Sci. USA* **8 6**:414-418

Dayton A.I., Sodroski J.G., Rosen C.A. *et al.*(1986) The trans-activator gene of the human T cell lymphotrophic virus type III is required for replication. *Cell* **4 4**:941-947

Delassus S., Cheynier R. & Wain-Hobson S. (1991) Evolution of human immunodeficiency virus type 1 nef and long terminal repeat sequences over 4 years in vivo and in vitro. *J. Virol.* **6 5**:225-231

Delassus S., Meyerhans A., Cheynier R. & Wain-Hobson S. (1992) Absence of selection of HIV-1 variants in vivo based on transcription/transactivation during progression to AIDS. *Virology* **1 8 8**:811-818

Descombes P. & Schibler U. (1991) A liver enriched transcriptional activator protein, LAP, and a transcriptional inhibitory protein, LIP, are transcribed from the same mRNA. *Cell* **6 7**:569-579

Diamond M.I., Miner J.N., Yoshinaga S.K. & Yamamoto K.R. (1990) Transcription factor interactions: selectors of positive or negative regulation from a single DNA element. *Science* **2 4 9**:1266-1272

Dingman J.D., Lebovitz R.M. & Roeder R.G. (1983) Accurate transcription inititation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nuc. Acid. Res.* **5**:1475-1489

Dollard S.C., Gummuluru S., Tsang S., Fultz P.N. & Dewhurst S. (1994) Enhanced responsiveness to nuclear factor NF-κB contributes to the unique phenotype of simian immunodeficiency virus variant $SIV_{smmPBj14}$. *J. Virol.* **6 8**:7800-7809

Donehower L.H. & Varmus H.E. (1984) A mutant murine leukemia virus with a single missense codon in pol is defective in a function affecting integration. *Proc. Natl. Acad. Sci. USA* **8 1**:6461-6465

Doucas V., Spyrou G. & Yaniv M. (1991) Unregulated expression of c-Jun or c-Fos proteins but not Jun D inhibits oestrogen receptor activity in human breast cancer cells. *Embo J.* **11**:2237-2245

Du W. & Maniatis T (1992) An ATF/CREB binding site is required for virus induction of the human interferon ß gene. *Proc. Natl. Acad. Sci. USA* **89**:2150-2154

Du W., Thanos D.& Maniatis T. (1993) Mechanisms of transcriptional synergism between distinct virus inducible enhancer elements. *Cell* **74**:887-898

Embretson J., Zupancic M., Ribas J.L. *et al.* (1993) Massive covert infection of helper T lymphocytes and macrophages by HIV during the incubation period of AIDS. Nature **362**:359-362

Englund G., Hoggan M.D., Theodore T.S., Martin M.A. (1991) A novel isolate containing alterations affecting the NF-kB element. *Virology* **181**:150-157

Faisst S. & Meyer S. (1992) Compilation of vertebrate encoded transcription factors. *Nuc. Acids Res.* **20**:3-26

Farina A.R., Davis-Smyth T., Gardner K. & Levens D. (1993) An early response of an AP-1-JunD complex during T-cell activation. *J. Biol. Chem.* **268**:26466-26475

Farr A., & Roman A. (1991) A pitfall of using a second plasmid to determine transfection efficiency. *Nuc. Acids Res.* **20**:920

Feinberg M.B., Baltimore D. & Frankel A.D. (1991) The role of Tat in the human immunodeficiency virus life cycle indicates a primary effect on transcription elomgation. *Proc. Natl. Acad. Sci. USA* **88**:4045-4049

Fisher A.G., Feinberg S.F., Josephs M.E. *et al.* (1986) The trans -activator gene of HTLV-III is essential for virus replication. *Nature* **320**:367-371

Fisher A.G., Ensoli B., Ivanoff L. *et al.* (1987) The sor gene product of HTLV-I is required for efficient virus transmission in vitro. *Science* **2 3 7**:888-893

Fodor E., Weinrich S.L., Meister A. *et al.* (1991) A pancreatic exocrine cell factor and AP-4 bind overlapping sites in the amalyse 2A enhancer. *Biochemistry* **3 0**:8102-8108

Foletta V.C., Sonobe M.H., Suzuki T. *et al.* (1994) Cloning and characterisation of the mouse fra-2 gene. *Oncogene* **9**:3305-3311

Franzoso G., Biswas P., Poli G. *et al.* (1994) A family of serine proteases expressed exclusively in myelo-monocytic cells specifically processes the nuclear factor-kB subunit p65 in vitro and may impair human immunodeficiency virus replication in these cells. *J. Exp. Med.* **1 8 0**:1445-1456

Fujisawa J.I., Toita M., Yoshimura T. & Yoshida M. (1991) The indirect association of human T-cell leukemia virus Tax protein with DNA results in transcriptional activation. *J. Virol.* **6 5**:4525-4528

Gabuzda D.H., Hess J.L., Small J.A. & Clements J.E (1989) Regulation of the Visna Virus long terminal repeat in macrophages involves cellular factors that bind sequences containing AP-1 sites *Mol. Cell. Biol.* **9**:2728-2733

Galien R., Emanoil-Ravier R. & Mercier G. (1994) Differential effects of c-Jun and CREB on cAMP response element activation by Ha-ras. *Oncogene* **9**:1101-1108

Garcia J.A., Wu F.K., Mitsuyasu R & Gaynor R.B. (1987) Interactions of cellular proteins involved in the transcriptional regulation of the human immunodeficiency virus. *EMBO* **6**:3761-3770

Garrel J. & Campuzano S. (1991) The Helix-Loop-Helix domain: a common motif for bristles, muscles and sex. *Bioessays* **1 3**:493-498

Garrett E.D., Tiley L.S. & Cullen B.R. (1991) Rev activates expression of the human immunodeficiency virus type 1 *vif* and *vpr* genes. *J. Virol.* **65**:1653-1657

Gatti M. (1992) Functional elements in *Drosophila melanogastor* heterochromatin. *Annu. Rev. Genet.* **26**:239-275

Gdovin S.L. & Clements J.E. (1992) Molecular mechanisms visna virus TAT: Identification of the targets for transcriptional activation and evidence for post-transcriptional effects. *Virology* **188**:438-450

Geballe A.P., Ventura P., Stowring L. & Haase A.T. (1985) Quantitative analysis of visna virus replication in vivo. *Virology* **141**:148-154

Gendelman H.E., Narayan O., Molineaux S. *et al.* (1985) Slow, persistent replication of lentiviruses: role of tissue macrophages and macrophage precursors in bone marrow. *Proc. Natl. Acad. Sci. USA* **82**:7086-7090

Gendelman H.E., Narayan O., Kennedy-Stoskopf S. *et al.* (1986) Tropism of sheep lentiviruses for monocytes: susceptibility to infection and virus gene expression increase during monocyte maturation to macrophages. *J. Virol.* **58**:67-74

Georgsson G. & Palsson P.A. (1971) The histopathology of maedi. *Vet Path.* **8**:63-80

Georgsson G., Houwers D.J., Palsson P.A. & Petursson G. (1989) Some comparitive aspects of visna and AIDS. pp82-98 *in* Modern Pathology of AIDS and other Lentiviral Infections Eds. Racz P., Haase A.T. & Gluckman J.C., Krager, Basel.

Geyer P.K., Green M.M. & Corces V.G. (1990) Tissue-specific transcriptional enhancers may act in trans on the gene located in the homologous chromosome: the molecular basis of transvection in Drosophila. *EMBO J.* **9**:2247-2256

Giese K., Cox J. & Grosschedl R. (1992) The HMG domain of lymphoid enhancer factor 1

bends DNA and facilitates assembly of functional nucleoprotein structures. *Cell* **6 9**:185-195

Giese K. & Grosschedl R. (1993) LEF-1 contains an activation domain that stimulates transcription only in a specific context of factor binding sites. *EMBO J.* **1 2**:4667-4676

Gill G. & Tijan R (1992) Eukaryotic coactivators associated with the TATA box binding potein. *Curr. Opin. Genet. Dev.*

Gill G., Pascal E., Tseng Z.H. & Tijan R. (1994) A glutamine rich hydrophobic patch in transcription factor Sp-1 contacts the dTAFII110 component of the Drosophila TFIID complex and mediates transcriptional activation. *Proc. Natl. Acad. Sci. USA* **9 1**:192-196

Golemis E.A., Speck N.A. & Hopkins N. (1990) Alignment of U3 region sequences of mammalian type C retroviruses: identification of highly conserved motifs and implications for enhancer design. *J. Virol.* **6 4**:534-542

Goodrich J.A., Hoey T., Thut C.J., Admon A. & Tijan R. (1993) Drosophila TAFII40 interacts with both a VP16 activation domain and the basal transcription factor TFIIB. *Cell* **7 5**:519-530

Goodrich J.A. & Tijan R. (1994) Transcription factors IIE and IIH and ATP hydrolysis direct promoter clearance by RNA polymerase II. *Cell* **7 7**:145-156

Goodwin G.H. (1990) The analysis of sequence specific DNA binding proteins in cell extracts. in Gel Electrophoresis of Nucleic Acids Eds Rickwood D.& Hames B.D. Oxford University Press.

Gourdou I., Mazarin V., Querat G., Sauze N. & Vigne R. (1989) The open reading frame S of visna virus genome is a *trans*-activating gene. *Virology* **1 7 1**:170-178

Grazia U., Felli M.P., Vacca A. *et al.* (1994) Positive and negative regulation of the composite octamer motif of the interleukin 2 enhancer by AP-1, OCT-2 and retinoic acid receptor. *J. Exp. Med.* **1 8 0**:1485-1497

Green I.R., Fiskerstrand C., Bertoni G. *et al.* (1993) Expression and characterisation of bioactive recombinant ovine TNF-alpha: some species specificity in cytotoxic response to TNF. *Cytokine* **5**:213-223

Greenblatt J., Xao H., Pearson A. *et al.* (1994) Protein-protein interactions that regulate transcription by RNA polymerase II. abstr. L-001, p.2. Kestone Symp. Basic Aspects of Transcription Suppl. 18C.

Grosschedl R., Giese K. & Pagel J. (1994) HMG domain proteins: architectural elements in the assembly of nucleoprotein structures. *Trends. Genet.* **10**:94-100

Gruda M.C., Kovary K., Metz R. & Bravo R. (1994) Regulation of Fra-1 and Fra-2 phosphorylation differs during the cell cycle of fibroblasts and phosphorylation in vitro by MAP kinases affects DNA binding activity. *Oncogene* **9**:2537-2547

Haase A.T. & Varmus (1973) Demonstration of a DNA provirus in the lytic growth of visna virus. *Nature New Biology* **245**:237-239

Haase A.T., Stowring L., Griffen D. & Price D. (1977) The slow persistent infection caused by visna virus: role of host restriction. *Science* **195**:175-177

Haase A.T., Stowring J.D., Harris B. *et al.* (1982) Visna virus DNA synthesis and the tempo of infection in vitro. *Virology* **119**:399-410

Haase A.T., Staskus K.A., Couch L. *et al.* (1994) Latency in lentiviral infections: analysis by polymerase chain reaction in situ. pp1-9, *In* Animal models of HIV and other retroviral infections, Racz P, Letvi N.L. & Gluckman J.C. (eds), Krager, Basel.

Hahn S. (1993) Structure (?) and function of acidic activators. *Cell* **72**:481-483

Hai T. & Curran T. (1991) Cross-family dimerisation of transcription factors Fos/Jun and ATF/CREB alters DNA binding specificity. *Proc. Natl. Acad. Sci. USA* **88**:3720-3724

Halazonetis T.D., georgopoulos K., Greenberg M.E. & Leder P. (1988) c-Jun dimerises with itself and with c-Fos, forming complexes of different DNA binding affinities. *Cell* **5 5**:917-924

Hanazawa S., Takashita A., Amano S. *et al.* (1993) Tumor necrosis factor-alpha induces expression of monocyte chemotractant JE via fos and jun genes in clonal osteoblastic MC3T3-E1 cells. *J. Biol. Chem.* **2 6 8**:9526-32

Harrich D.J., Garcia R., Mitsuyasu R. & Gaynor R. (1990) TAR independent activation of the human immunodeficiency virus in phorbol ester stimulated T lymphocytes. *EMBO J.* **9**:4417-4423

Harris J.D., Scott J.V., Traynor B. *et al.* (1981) Visna virus DNA: discovery of a novel gapped structure *Virology* **11 3**:573-583

Harris J.D., Blum H., Scott J. *et al.* (1984) Slow virus visna: Reproduction in vitro of virus from extrachromosomal DNA. *Proc. Natl. Acad. Sci.* **8 1**:7212-7215

Hass R., Pfannkuche H.J., Kharbanda S. *et al.* (1991) Protein kinase C activation and protooncogene expression in differentiaition/retrodifferentiaition of human U-937 leukemia cells. *Cell Growth Differ.* **2**:541-8

Hauber J. & Cullen B.R (1988) Mutational analysis of the trans- activation responsive region of the human immunodeficiency virus type I long terminal repeat. *J. Virol.* **6 2**:673-679

Herschman H.R. (1991) Primary response genes induced by growth factors and tumor promoters. *Annu. Rev. Biochem.* **6 0**:281-319

Hess J.L., Clements J.E. & Narayan O. (1985) Cis- and trans-acting transcriptional regulation of visna virus. *Science* **2 2 9**:482-485

Hess J.L., Small J.A. & Clements J.E. (1989) Sequences in the Visna Virus long terminal repeat that control transcriptional activity and respond to viral transactivation: involvement of

AP-1 sites in basal activity and transactivation. *J. Virol.* **6 3**:3001-3010

Hidaka M., Inoue J., Yoshida M. & Seiki M. (1988) Post transcriptional regulator (rex) of HTLV-I initiates expression of viral structural proteins but supresses expression of regulatory. *EMBO* **7**:519-523

Hilberg F., Aguzzi A., Howells N. & Wagner E.F. (1993) c-Jun is essential for normal mouse development and hepatogenesis. *Nature* **3 6 5**:179-181

Hippenmeyer P.J. & Grandgenett D.P. (1984) Requirement of the avian retrovirus pp32 DNA binding protein domain for replication. *Virology* **1 3 7**:358-370

Hisatake K., Roeder R.G. & Horikoshi M. (1993) Functional dissection of TFIIB domains required for TFIIB-TFIID promoter complex formation and basal transcription activity. *Nature* **3 6 3**:744-747

Ho I.C. & Leiden J.M. (1990) Regulation of the human T cell receptor α gene enhancer: Multiple ubiquitous and T cell specific nuclear proteins interact with four hypomethylated enhancer elements. *Mol. Cell Biol.* **1 0**:4720-4727

Ho I.C., Yang L.H., Morle G. & Leiden J.M. (1989) A T cell specific transcriptional enhancer element 3' of the Cα in the human T-cell α locus. *Proc. Natl. Acad. Sci. USA* **8 6**:6714-6718

Hoey T., Weinzierl R.O.J., Gill G., *et al.* (1993) Molecular cloningand functional analysis of Drosophila TAF110 reveal properties expected of coactivators. *Cell* **7 2**:247-260

Hope I.A. & Struhl K. (1986) Functional dissection of a eukaryotic transcriptional activator protein, GCN4 of yeast. *Cell* **4 6**:885-894

Hope T.J., McDonald D., Huang X. *et al.* (1990) Mutational analysis of the human immunodefficiency virus type 1 Rev transactivator: essential residues near the amino terminus. *J. Virol.* **6 4**:5360-5366

Howell R.M., Fitzgibbon J.E., Noe M. *et al.* (1991) In vivo sequence variation of the human immunodeficiency virus type I *env* gene: evidence for recombination among variants found in a single individual. *AIDS Res. Human Retroviruses* **7**:869-876

Hsu W., Kerppola T.K., Chen P-L., Curran T. & Chen-Kiang S. (1994) Fos and Jun repress transcription activation by NF-IL6 through association at the basic zipper region. *Mol. Cell. Biol.* **14**:268-276

Hu K.Q., Vierling J.M. & Siddiqui A. (1990) Trans-activation of HLA-DR gene by hepatitis B virus X gene product. *Proc. Natl. Acad. Sci. USA* **87**:7140-7144

Huang L-M. & Jeang K-T. (1993) Increased spacing between Sp1 and TATAA renders human immunodeficiency virus type 1 replication defective: implications for Tat function. *J. Virol.* **67**:6937-6944

Inostroza J.A., Mermelstein F.H., Ha I. et al (1992) Dr1, a TATA binding ptotein associated phosphoprotein and inhibitor of class II gene transcription. *Cell* **70**:477-489

Ivashkiv L.B., Liou H-C., Kara C.J. (1990) mXBP/CRE-BP2 and c-Jun form a complex which binds to the cyclic AMP, but not to the 12-O-tetradecanoylphorbol-13-acetate, response element. *Mol. Cell Biol.* **10**:1609-1621

Jeang K-T., Berkhout B. & Dropulic B. Effects of integration and replication on transcription of the HIV long terminal repeat. *J. Biol. Chem.* **268**:24940-24949

Johnson P.F. & McKnight S.L. (1989) Eukaryotic transcriptional regulatory proteins. *Annu. Rev. Biochem.* **58**:799-839

Jones N. (1990) Transcriptional regulation by dimerisation: two sides to an incestuous relationship. *Cell* **61**:9-11

Jones K.A. (1993) Tat and the HIV-1 promoter. *Cur. Opin. Cell Biol.* **5**:461-468

Jones K.A. & Peterlin B.M. (1994) Control of RNA initiation and elongation at the HIV-1 promoter. *Annu. Rev. Biochem.* **6 3**:717-743

Jain J., Valge-Archer V.E. & Rao A. (1992) Analysis of the AP-1 sites in the IL-2 promoter. *Proc. Natl. Acad. Sci. USA* **1 4 8**:1240-1250

Ju. G. & Cullen B.R. (1985) The role of avian retroviral LTRs in the regulation of gene expression and viral replication. *Adv. Virus Res.* **3 0**:179-223

Juan L-J., Utley R.T., Adams C.C. *et al.* (1994) Differential repression of transcription factor binding by histone H1 is regulated by the core histone amino termini. *EMBO J.* **1 3**:6031-6040

Judd B.H. (1988) Transvection: allelic cross talk. *Cell* **5 3**:841-843

Junghans R.P., Boone L.R. & Skalka A.M. (1982) Products of reverse transcription in avian retroviruses analyzed by electron microscopy. *J. Virol.* **4 3**:544-554

Kamakaka R.T. & Thomas J.O. (1990) Chromatin structure of transcriptionally competent and repressed genes. *EMBO J.* **9**:3997-4006

Kamine J., Subramanian T., Chinnaduria G. (1991)Sp-1 dependent activation of a synthetic promter by human immunodeficiency virus type 1 Tat protein. *Proc. Natl. Acad. Sci. USA* **8**:8510-8514

Kanno T., Brown K., Franzoso G. & Siebenlist U. (1994) Kinetic analysis of human T-cell leukemia virus type 1 Tax mediated activation of NF-kB. *Mol. Cell Biol.* **1 4**:6443-6451

Karin M. & Smeal T. (1992) Control of transcription factors by signal transduction pathways: the beginning of the end. *Trends in Biol. Sci.* **1 7**:418-422

Kashanchi F., Shibata R., Ross E.K., Brady J.N. & Martin M.A. (1994) Second-site long terminal repeat (LTR) revertants of replication-defective human immunodeficiency virus:

effects of revertatnt TATA box motifs on virus infectivity LTR-directed expression, in vitro RNA synthesis, and binding of basal transcription factors TFIID and TFIIA. *J. Virol.* **6 8**:3298-3307

Katz R.A. & Skalka A.M. (1990) Generation of diversity in retroviruses. *Annu. Rev. Genet.* **2 4**:409-445

Katzman M. & Sudol M. (1994) In vitro activities of purified visna virus integrase. *J. Virol.* **6 8**:3558-3569

Kawakami T., Sherman L., Dahlberg J. et al (1987) Nucleotide sequence analysis of Equine Infectious Anemia Virus proviral DNA. *Virol.* **1 5 8**:300-312

Kerppola T.K. & Curran T. (1991a) Fos-Jun heterodimers and Jun homodimers bend DNA in opposite orientations: implications for transcription factor cooperativity. *Cell* **6 6**:317-326

Kerppola T.K. & Curran T. (1991b) DNA bending by Fos and Jun: the flexible hinge model. *Science* **2 5 4**:1210-1214

Kessler M., & Mathews M.B. (1992) Premature termination and processing of human immunodeficiency virus type-1 promoted transcripts. *J.Virol.* **6 6**:4488-4496

Khalili K., Khoury G. & Brady J. (1986) Spacing between simian virus 40 early transcriptional control sequences is important for regulation of early RNA synthesis and gene expression. *J.Virol.* **6 0**:935-942

Kharroubi A.E. & Verdin E. (1994) Protein-DNA interactions within DNAse I-hypersensitive sites located downstream of the HIV-1 promoter. *J. Biol. Chem.* **2 6 9**:19916-19924

Kim J.Y.H., Gonzalez-Scarano F., Zeichner S.L. & Alwine J.C. (1993) Replication of type-1 human immunodeficiency viruses containing linker substitution mutations in the -210 to -130 region of the long terminal repeat. *J. Virol.* **6 7**:1658-1662

Klampfer L., Lee T.H., Hsu W., Vilcek J. & Chen-Kiang S. (1994) NF-IL-6 and AP-1 cooperatively modulate the activation of the TSG-6 gene by tumour necrosis factor alpha and interleukin-1. *Mol. Cell. Biol.* **1 4**:6561-6569

Klein C. & Struhl K (1994) Increased recruitment of TATA binding protein to the promoter by transcriptional activation domains in vivo. *Science* **2 6 6**:280-282

Klaver B. & Berkhout B. (1994) Comparison of 5' and 3' long terminal repeat promoter function in human immunodeficiency virus. *J. Virol.* **6 8**:3830-3840

Knight S.C., Macatonia S.E. & Patterson S. (1992) Antigen presentation in HIV infection. *in* Immunology of HIV Infection, Ed. Bird A.G., Immunology & Medecine Series, Kluwer Academic Publishers.

Koken S.E.C., Wamel J.L.B., Goudsmit J., Berkhout B. & Geelen J.L.M.C. (1992) Natural variants of the HIV-1 long terminal repeat: analysis of promoters with duplicated DNA regulatory motifs. *Virology* **1 9 1**:968-972

Kolson D.L. Buchhalter J., Collman R. *et al.* (1993) HIV-1 Tat alters normal organization of neurons and astrocytes in primary rodent brain cell cultures: RGD sequence dependence. *AIDS Res. Hum. Retro.* **9**:677-685

Kovary K. & Bravo R. (1991) Expression of different Jun and Fos proteins during the Go to G1 transition in mouse fibroblasts: in vitro and in vivo associations. *Mol. Cell. Biol.* **11**:2451-2459

Kriste T.M., Sharp P.A. (1990) Interactions of the Oct-1 POU sub-domains with specific DNA sequences and with the HSV α-*trans* activator protein. *Genes Dev.* **1 2**:2383-2396

LaFemina R.L., Schneider C.L., Robbins H.L. *et al.* (1992) Requirement for active human immodeficiency virus type 1 integrase enzyme for productive infection of human T-lymphoid cells. *J. Virol.* **6 6**:7414-7419

Lairmore M.D., Rosadio R.H. & De Martini J.C. (1986) Ovine lentivirus lymphoid interstitial pneumonia. Rapid induction in neonatal lambs. *Am. J. Pathol.* **1 2 5**:173-181

Lamb P. & McKnight S.L. (1991) Diversity and specificity in transcriptional regulation: the benefits of heterotypic dimerisation. *Trends Biol. Sci.* **1 6**:417-422

Lane D., Prentki P. & Chandler M. (1992) Use of gel retardation to analyze protein-nucleic acid interactions. *Microbiol. Rev.* **5 6**:509-528

Laspia M.F., Rice A.P., Mathews M.B. (1989) HIV-1 Tat protein increases transcriptional initiation and stabilises elongation. *Cell* **5 9**:283-292

LeClair K.P., Blanar M.A. & Sharp P.A. (1992) The p50 subunit of NF-κB associates with the NF-IL6 transcription factor. *Proc. Natl. Acad. Sci. USA* **8 9**:8145-8149

Leonard J.C., Parott A.J., Buckler-White W. *et al.* (1989) The NF-kB binding sites in the human immunodeficiency virus type 1 long terminal repeat are not required for virus infectivity. *J. Virol.* **6 3**:4919-4924

Leung K. & Nabel G.J. (1988) HTLV-1 transactivator induces interleukin-2 receptor expression through an NF-κB like factor. *Nature* **3 3 3**:776-778

Lewin B. (1994) Chromatin and gene expression:constant questions, but changing answers. *Cell* **7 9**:397-406

Li Q. & Wrange O. (1993) Translational positioning of a nucleosomal glucocorticoid response element modulates glucocorticoid receptor affinity. *Genes Dev.* **7**:2471-2482

Lin F.H. & Thormar H. (1970) Ribonucleic acid-dependent doexyribonucleic acid polymerase in visna virus. *J. Virol.* **6**:702-704

Lin Y-S. & Green M.R. (1991) Mechanism of action of an acidic transcriptional activator in vitro.

Cell **6 4**:971-981

Lord K.A., Abdollahi A., Hoffman-Liebermann B. & Lieberman D.A. (1993) Proto-oncogene of the Fos/Jun family of transcription factors are positive regulators of myeloid differentiation. *Mol. Cell. Biol.* **1 3**:841-851

Maguire H.F., Hoeffler J.P. & Siddiqui A. (1991) HBV X protein alters the DNA binding specificity of CREB and ATF-2 by protein-protein interactions.

Majello B., Luca P.D., Hagen G., Suske G. & Lania L. (1994) Different members of the Sp1 multigene family exert opposite transcriptional regulation of the long terminal repeat of HIV-1. *Nuc. Acids Res.* **2 2**:4914-4921

Malim M.H., Bohnlein S., Hauber J. & Cullen B.R. (1989) Functional dissection of the HIV-1 rev transactivator - derivation of a trans dominant repressor of Rev function. *Cell* **5 8**:205-214

Malim M.H. & Cullen B.R. (1991) HIV-I structural gene expression requires the binding of multiple Rev monomers to the viral RRE: implications for HIV-I latency. *Cell* **6 5**:241-248

Matsui M., Tokuhara M., Konuma Y., Nomura N. & Ishizaki R. (1990) Isolation of human fos-related genes and their expression during monocyte-macrophage differentiation. *Oncogene* **5**:249-255

Maury W. (1994) Monocyte maturation controls expression of equine infectious anemia virus. *J. Virol.* **6 8**:6270-6279

McPherson C.E., Shim E-Y., Friedman D.S. & Zaret K.S. (1993) An active tissue specific enhancer and bound transcription factors existing in a precisely positioned nucleosome array. *Cell* **7 5**:387-398

Mermod N., Williams T.J. & Tijan R. (1988) Enhancer binding factors AP-4 and AP-1 act in concert to activate SV40 late transcription in vitro. *Nature* **3 3 2**:557-561

Mermod N., O'Neill E.A., Kelly T.J. & Tijan R. (1989) The proline rich transcriptional activatorof CTF/NF-1 is distinct from the replication and DNA binding domain. *Cell* **5 8**:741-753

Metz R. & Ziff E. (1991) cAMP stimulates the C/EBP related transcription factor rNF-IL6 to translocate to the nucleus and induce *c-fos* transcription. *Genes Dev.* **5**:1754-1766

Meyerhans A., Cheynier R., Albert J. *et al.* (1989) Temporal fluctuations in HIV quasispecies in vivo are not reflected by sequential HIV isolations. *Cell* **5 8**:901-910

Michael N.L., D'Arcy L., Ehrenberg P.K. & Redfield R.R. (1994) Naturally ocurring genotypes of the human immunodeficiency virus type 1 long terminal repeat display a wide range of basal and Tat-induced transcriptional activities. *J. Virol.* **6 8**:3163-3174

Miner J.N. & Yamamoto K.R. (1992) The basic region of AP-1 specifies glucocorticoid receptor activity at a composite response element. *Genes and Dev.* **6**:2491-2501

Miyazawa T., Kohmoto M., Kawaguchi Y. *et al.* (1993) The AP-1 binding site in the feline immunodeficiency virus long terminal repeat is not required for virus replication in feline T lymphocytes. *J. Gen. Virol.* **7 4**:1573-1580

Miyazawa T., Tomonaga K., Kawaguchi Y. & Mikami T. (1994a) The genome of feline immunodeficiency virus. *Arch. Virol.* **1 3 4**:221-234

Miyazawa T. Tomonaga K., Kawaguchi Y. *et al.* (1994) Effects of insertion of multiple AP-1 binding sites into the U3 region of the long terminal repeat of feline immunodeficiency virus type 1. *Arch. Virol.* **1 3 9**:37-38

Monteclaro F.S. & Vogt P.K. (1993) A Jun-binding protein related to a putative tumour suppressor . *Proc. Natl. Acad. Sci. USA* **9 0**:6726-6730

Muegge K., Vila M., Gusella G.L. *et al.* (1993) Interleukin 1 induction of the c-jun promoter. *Proc. Natl. Acad. Sci. USA* **9 0**:7054-7058

Mueller P.R. & Wold B. (1989) In vivo footprinting of a muscle specific enhancer by ligation mediated PCR. *Science* **2 4 6**:780-786

Muesing M.A., Smith D.H. & Capon D.J. (1987) Regulation of mRNA accumulation by a human immunodeficiency virus trans -activator protein. *Cell* **4 8**:691-701

Muller H.P., & Varmus H.E. (1994) DNA bending creates favoured sites for retroviral integration: an explanation for preferred integration sites in nucleosomes. *EMBO J.* **1 3**:4704-4714

Munoz E., Courtois G., Veschambre P. Jalinot P. & Israel A. (1994) Tax induces nuclear translocation of NF-κB through dissociation of cytoplasmic complexes containing p105 or p100 but does not induce degradation of IκBα/MAD3. *J. Virol.* **6 8**:8035-8044

Nacheva B.A., Guschin D.Y., Preobrazhenskaya O.V. *et al.* (1989) Change in the pattern of histone binding to DNA upon transcriptional activation. *Cell* **5 8**:27-36

Nagashunmugan T., Velpandi A., Otsuka T. *et al.* (1992) Analysis of the viral determinants underlying replication kinetics and cellular tropism of human immunodeficiency virus. *Pathobiology* **6 0**:234-245

Nakabeppu Y., Ryder K. & Nathans D. (1988) DNA binding activities of three murine Jun proteins: stimulation by Fos. *Cell* **5 5**:907-915

Narayan O., Wolinsky J.S., Clements J.E. *et al.* (1982) Slow virus replication: the role of macrophages in the persistence and expression of visna viruses of sheep and goats. *J. Gen. Virol.* **5 9**:345-356

Narayan O., Kennedy-Stoskopf S., Sheffer D. *et al.* (1983) Activation of caprine arthritis-encephalitis virus expression during maturation of monocytes to macrophages. *Infect. Immun.* **4 1**:67-73

Natesan S. & Gilman M.Z. (1993) DNA bending and orientation-dependent function of YY1 in the c-fos promoter. *Genes Dev.* **7**:2497-2509

Neuveut N.,Vigne N., Clements J.E., & Sire J. (1993) The Visna transcriptional activator TAT: Effects on the viral LTR and on cellular genes. *Virology* **197**:236-244

Ng L., Forrest D. & Curran T. (1993) Differential roles for Fos and Jun in DNA-binding: redox-dependent and independent functions. *Nucleic Acids Res.* **21**:5831-5837

Nichols M.F. Weih F., Schmid W. et al., (1992) Phosphorylation of CREB affects its binding to high and low affinity sites: implications for cAMP induced gene transcription. *EMBO J.* **11**:3337-3346

Nielsen A.E., Pallisgaard N., Pedersen F.S. & Jorgensen (1994) Basic helix-loop-helix proteins in murine type C retrovirus transcriptional regulation. *J. Virol.* **68**:5638-5674

Nixon D.F. (1992) The cytotoxix T cell response to HIV. *in* Immunology of HIV Infection, Ed. Bird A.G., Immunology & Medecine Series, Kluwer Academic Publishers.

Noll M. & Kornberg R.D. (1977) Action of micrococcal nuclease on chromatin and the location of histone H1. *J. Mol. Biol.* 109:393-404

Nyborg J.K., Dynan W.S., Chen I.S.Y. & Wachsman W. (1988) Binding of host cell factors to DNA sequences in the long terminal repeats of human T-cell leukemia virus type I: implications for viral gene expression. *Proc. Narl. Acad. Sci. USA* **85**:1457-1461

Ofir R., Dwarki V.J., Rashud D. & Verma I.M. (1990) Phosphorylation of the C-terminus of Fos protein is required for transcriptional repression of the c-fos promoter. *Nature* **348**:80-82

Oliver R.E., Gorham J.R., Parish S.F., Hadlow W.J. & Narayan O. (1981) Ovine progressive pneumonia: pathological and virological studies on the naturally occuring disease. *Am J. Vet. Res.* **42**:1554-1559

Oliviero S. & Struhl K. (1991) Synergistic transcriptional enhancement does not depend on the number of acidic activation domains bound to the promoter. *Proc.Natl. Acad. Sci. USA* **8 8**:224-228

Olmsted R.A., Hirsch V.M., Purcell R.H. & Johnson P.R. (1989a) Nucleotide sequence analysis of feline immunodeficiency virus: genome organization and relationship to other lentiviruses. *Proc. Natl. Acad. Sci. USA* **8 6**:8088-8092

Olmstead R.A., Barnes A.K., Yamamoto J.K. et al (1989b) Molecular cloning of feline immunodeficiency virus. *Proc. Natl. Acad. Sci. USA* **8 6**:2448-2452

Ondek B., Gloss L. & Herr W. (1988) The SV40 enhancer contains two distinct levels of organisation. *Nature* **3 3 3**:40-45

Panganiban A.T. & Temin H.M. (1984) Circles with two tandem LTRs are precursors to integrated retrovirus DNA. *Cell* **3 6**:673-679

Pantaleo G., Graziosi C., Demarest J.F. *et al.* (1993) HIV infection is active and progressive in lymphoid tissue during the clinically latent stage of disease. *Nature* **3 6 2**:355-358

Papivassilou A.G., Treier M. *et al.* (1992) Targeted degradation of c-fos, but not v-fos, by a phosphorylation dependent signal on c-Jun. *Science* **2 5 8**:1941-1944

Peluso R., Haase A., Stowring L. *et al.* (1985) A trojan horse mechanism for the spread of visna virus in monocytes. *Virology* **1 4 7**:231-236

Pennings S., Meersseman G., Bradbury E.M. (1994) Linker histones H1 and H5 prevent the mobility of positioned nucleosomes. *Proc. Natl. Acad. Sci. USA* **9 1**:10275-10279

Perkins N.D., Edwards N.L., Duckett C.S. *et al.* (1993) A cooperative interaction between NF-kB and Sp-1 is required for HIV-1 enhancer activation. *EMBO J.* **1 2**:3551-3558

Perkins N.D., Agranoff A.B., Pascal E. & Nabel G.J. (1994) An interaction between the DNA-binding domain of Rel A(p65) and Sp1 mediates human immunodeficiency virus gene activation. *Mol. Cell. Biol.* **1 4**:6570-6583

Perrin S. & Gilliland G. (1990) Site-specific mutagenesis using asymmetric polymerase chain reaction and a single mutant primer. *Nucl. Acids Res.* **1 8**:7433-7438

Pfeifer G.P. & Riggs A.D. (1991) Chromatin differences between active and inactive X chromosomes revealed by genomic footprinting of permeabilized cells using DNase I and ligation-mediated PCR. *Genes Dev.* **5**:1102-1113

Philippon V., Vellutini C., Gambarelli D. *et al.* (1994) The basic domain of lentiviral Tat protein is responsible for damages in mouse brain: involvement of cytokines. *Virology* **2 0 5**:519-529

Pierrou S., Hellqvist M., Samuelsson L., Enerback S. & Carlsson P. (1994) Cloning and charatcerisation of seven human forkhead proteins: binding site specificity and DNA bending. *EMBO J.* **1 3**:5002-5012

Pomerantz R.J., Trono D., Feinberg M.B. & Baltimore D. (1990) Cells nonproductively infected with HIV-1 exhibit an aberrant pattern of viral RNA expression: a molecular model for latency. *Cell* **6 1**:1271-1276

Pruss D., Reeves R., Bushman F.D. & Wolfe A.P. (1994) The influence of DNA and nucleosome structure on integration events directed by HIV integrase. *J. Biol. Chem.* **2 6 9**:25031-25041

Pugh B.F. & Tijan R. (1990) Mechanism of transcriptional activation by Sp-1: evidence for coactivators. *Cell* **6 1**:1187-1197

Pulverer B.J., Kyriakis J.M., Avruch J. (1991) Phosphorylation of c-jun mediated by MAP kinases. *Nature* **3 5 3**:670-672

Querat G., Audoly G., Sonigo P., Vigne R. (1990) Nucleotide sequence analysis of SA-OMVV, a visna related lentivirus: phylogenetic history of lentiviruses. *Virology* **1 7 5**:434-447

Rahuel C., Vinit M-A., Lemarchandel V., Cartron J-P. & Romeo P-H. (1992) Erythroid-specific activity of the glycophorin B promoter requires GATA-1 mediated displacement of a repressor. *EMBO J.* **11**:4095-4102

Rao A. (1994) NF-ATp: atranscription factor required for the co-ordinate induction of several cytokine genes. *Immunol. Tod.* **1 5**:274-281

Redner R.L. Lee A. W-M., Osawa G.A. & Nienhuis A.W. (1992) variable pattern of jun and fos gene expression in different hematopoietic cell lines during interleukin 3-induced entry into the cell cycle. *Oncogene* **7**:43-50

Rimsky L., Hauber J., Dukovich M. *et al.* (1988) Functional replacement of the HIV rev protein by the HTLV-I rex protein. *Nature* **3 5**:738-740

Roberts S.G.E., Ha I., Maldonado E., Reinberg D. & Green M.R. (1993) Interaction between an acidic activator and transcription factor TFIIB is required for transcriptional activation. *Nature* **3 6 3**:741-744

Roberts S.G.E., Green M.R. (1994) Activator induced conformational change in general transcription factor TFIIB. *Nature* **3 7 1**:717-720

Rosen C.A., Sodroski J.G. & Haseltine W.A. (1985) The location of cis-acting regulatory sequences in the human T cell lymphotropic virus type III (HTLV-III/LAV) long terminal repeat. *Cell* **4 1**:813-823

Ross E.K., Buckler-White A.J., Rabson A.B. *et al.* (1991) Contribution of NF-kB and Sp-1 binding motifs to the replicative capacity of human immunodeficiency virus type -1: distinct patterns of viral growth are determined by T-cell types. *J. Virol.* **6 5**:4350-4358

Roth M.J., Schwartzberg P., Tanese N. & Goff S.P. (1990) Analysis of mutations in the integration function of Moloney murine leukemia virus, effects on DNA binding and cutting. *J. Virol.* **6 4**:4709-4717

Roux P., Blanchard J.-M., Fernandez A. *et al.* (1990) Nuclear localization of c-Fos, but not v-Fos proteins, is controlled by extracellular signals. *Cell* **6 3**:341-351

Ryder K., Lau L.F. & Nathans D. (1988) A gene activated by growth factors is related to the oncogen v-jun. *Proc. Natl. Acad. Sci. USA* **8 5**:1487-1491

Ryder K. & Nathans D. (1988) Induction of proto-oncogene c-jun by serum growth factors. *Proc. Natl. Acad. Sci. USA* **8 5**:8464-8467

Ryder K., Lanahan A., Perez-Albuerne E. & Nathans D. (1989) JunD: a third member of the jun gene family. *Proc. Natl. Acad. Sci. USA* **8 6**:1500-1503

RyseckR.P., Hirai S.I., Yaniv M. Bravo R. (1988) Transcriptional activation of c-jun during Go/G1 transition in mouse fibroblasts. *Nature* **3 3 4**:535-537

Ryseck R.P. & Bravo R. (1991) c-Jun, JunB and JunD differ in their binding affinities to AP-1 and CRE consensus sequences: effects of Fos proteins. *Oncogene* **6**:533-542

Saffen D.W., Cole A.J., Worley P.F. *et al.* (1988) Convulsant-induced increase in transcription factor messenger RNAs in rat brain. *Proc. Natl. Acad. Sci. USA* **8 5**:7795-7799

Sassone-Corsi P., Ransone L.J., Lamph W.W. & Verma I.M. (1988) Direct interaction between fos and jun nuclear oncoproteins: role of the 'leucine zipper' domain. *Nature* **3 3 6**:692-695

Sakai H., Kawamura M., Sakuragi J-I. *et al.* (1993) Integration is essential for efficient gene expression of human immunodeficiency virus type 1. *J. Virol.* **6 7**:1169-1174

Saltarelli M., Querat G., Konings D.A.M. *et al.* (1990) Nucleotide sequence analysis of

molecular clones of CAEV which generate infectious virus. *Virology* **179**:347-364

Saltarelli M., Schoborg R., Gdovin S.L. & Clements J.E. (1993) Tha CAEV *tat* gene *trans*-activates the viral LTR and is necessary for efficient viral replication. *Virol.* **197**:35-44

Sambrook J., Fritsch E.F., & Maniatis T. (1989) In Molecular Cloning: A Laboratory Manual. 2nd edition Cold Spring Harbor Laboratory press, NY.

Sandmeyer S.B., Hansen L.J. & Chalker D.L. (1990) Integration specificity of retrotransposons and retroviruses. *Annu. Rev. Genet.* **24**:491-518

Sargan D.R. & Bennet I.D. (1989) A transcriptional map of visna virus: definition of the second intron structure suggests a rev-like gene product. *J. Gen. Virol.* **70**:1995-2006

Sargan D.R., Bennet I.D., Cousens C. *et al.* (1991) Nucleotide sequence of EV-1, a British isolate of maedi-visna virus. *J. Gen. Virol.* **72**:1893-1903

Sargan D.R., Sutton K.A., Bennet I., McConnel I., Harkiss G. (1995) LTR variants in Maedi Visna virus. *Virology* In press

Satchwell S.C., Drew H.R. & Travers A.A. (1986) Periodicities in chicken nucleosome core DNA *J. Mol. Biol.* **191**:659-675

Satchwell S.C. & Travers A.A. (1989) Asymmetry and polarity of nucleosomes in chicken erythrocyte chromatin. *EMBO J.* **8**:229-238

Sattentau Q.J. & Weiss R.A. (1988) The CD4 antigen: physiological ligand and HIV-1 receptor. *Cell* **52**:631-633

Schatz C. & Chatton B. (1990) Stereoalignment requirements for activation of transcription by the Simian virus 40 enhancer. *Nuc. Acids Res.* **18**:421-427

Schleif R. (1992) DNA looping. *Annu. Rev. Biochem.* **6 1**:199-223

Schreiber E., Matthias P., Muller M.M. & Schaffner W. (1989) Rapid detection of octamer binding proteins with 'mini extracts', prepared from a small number of cells. *Nuc. Acids Res.* **1 7**:6419

Schrier R.D., Gnann J.W., Landes R. *et al*. (1989) T cell recognition of HIV synthetic peptides in a natural infection. *J. Virol.* **1 4 2**:1166-1176

Schuitmaker H., Groenink M., Meyaard L. *et al.* (1993)Early replication steps but not cell type specific signalling of the viral long terminal repeat determine HIV-1 monocytotropism. *AIDS Res. Hum. Retro.* **9**:669-675

Schutte J., Viallet J., Nau M. *et al.* (1989) junB inhibits and c-fos stimulates the transforming and trans-activating activities of c-jun. *Cell* **5 9**:987-997

Schwedler U., Song J., Aiken C. & Trono D. (1993) Vif is crucial for human immunodeficiency virus type 1 proviral DNA synthesis in infected cells. *J. Virol.* **6 7**:4945-4955

Seed B. & Sheen J-Y. (1988) A simple phase extraction assay for chloramphenicol acetyltransferase activity. *Gene* **6 7**:271-277

Segall A.M., Goodman S.D. & Nash H.A. (1994) Architectural elements in nucleoprotein complexes: interchangeability of specific and non-specific elements. *EMBO J.* **1 3**:4536-4548

Seipel K., Georgiev O. & Schaffner W. (1992) Different activation domains stimulate transcription from remote ('enhancer') and proximal ('promoter') positions. *EMBO* **11**:4961-4968

Semmes O.J. & Jeang K.T. (1992) Mutational analysis of human T-cell leukemia virus type I Tax: regions necessary for function determined with 47 mutant proteins. *J. Virol.* **6 6**:7183-7192

Seto E., Mitchell P.J. & Yen T.S.B. (1990) Transactivation by the hepatitis B virus X protein depends on AP-2 and other transcription factors. *Nature* **2 4 4**:73-74

Sharpless N.E., O'Brien W.A., Verdin E. *et al.* (1992) Human immunodeficiency virus type 1 tropism for brain microglial cells is determined by a region of the env glycoprotein that also controls macrophage tropism. *J. Virol.* **6 6**:2588-2593

Shih C-C., Stoye J.P. & Coffin J.M. (1988) Highly preferred targets for retroviral integration. *Cell* **5 3**:531-537

Shih D.S., Carruth L.M., Anderson M. & Clements J.E. (1992) Involvment of Fos and Jun in the activation of visna virus gene expression in macrophages through a AP-1 site in the viral LTR. *Virology* **1 9 0**:84-91

Shin. C.-G., Taddeo W.A., Haseltine W.A. & Farnet C.M. (1994) genetic analysis of the human immunodeficiency virus type 1 integrase protein. *J. Virol.* **6 8**:1633-1642

Shioda T., Levy J.A. & Cheng-Mayer C. (1991) Macrophage and T-cell line tropisms of HIV-1 are determined by specific regions of the gp120 gene. *Nature* **3 4 9**:167-169

Siddiqui A., Gaynor R.,Srinivasan A. *et al.* (1989) Trans-activation of viral enhancers including long terminal repeat of the human immuno deficiency virus by the Hepatitis virus X protein. *Virology* **1 6 9**:479-484

Sigurdsson B. (1954a) Maedi, a slow progressive pneumonia of sheep: an epizoological and a pathological study. *Brit. Vet. J.* **11 0**:255-270

Sigurdsson B (1954b) Rida, a chronic encephalitis of sheep. *Brit. Vet. J.* **11 0**:341-354

Sigurdsson B., Palsson P.A. & Grimsson H. (1957) Visna, a demyelinating transmissible disease of sheep. *J. Neuropathol. Exp. Neurol.* **1 6**:389-403

Sigurdsson B. & Palsson P.A. (1958) Visna of sheep. A slow demyelinating infection. *Brit. J. Exp. Pathol.* **3 9**:519-528

Simpson R.T. (1978) Structure of the chromatosome, a chromatin particle containing 160 base pairs of DNA and all the histones.*Biochemistry* **1 7**:5524-5531

Smale S.T., Schmidt M.C., Berk A.J. & Baltimore D. (1990) Transcriptional activation by Sp-1 as directed through TATA or initiator: specific requirement for mammalian transcription factor TFIID. Proc. *Natl. Acad. Sci. USA* **8 7**:4509-4513

Small J.A., Bieberich C., Ghotbi Z. *et al.* (1989) The visna virus long terminal repeat directs expression of a reporter gene in activated macrophages, lymphocytes and the central nervous systems of transgenic mice. *J. Virol.* **6 3**:1891-1896

Smith M.R. & Greene W.C. (1991) Identification of HTLV-I Tax transactivator mutants exhibiting nivel transcriptional phenotypes. *Genes Dev.* **4** :1875-1885

Solomon M.J., Larsen P.L. & Varshavsky (1988) Mapping protein-DNA interactions in vivo with formaldehyde: evidence that histone H4 is retained on a highly transcribed gene. *Cell* **5 3**:937-947

Sonigo P., Alizon M., Staskus K. *et al.* (1985) Nucleotide sequence of the visna lentivirus: relationship to the AIDS virus. *Cell* **4 2**:369-382

Southgate C.D. & Green M.R. (1991) The HIV-1 Tat protein activates transcription from an upstream DNA-binding site: implications for Tat function. *Genes Dev.* **5**:2496-2507

StaskusK.A., Retzel E.F., Lewis E.D., *et al.* (1991) Isolation of replication competent molecular clones of visna virus. *Virology* **1 8 1**:228-240

Stein B., Baldwin A.S., Ballard D.W. *et al.* (1993) Cross-coupling of the NF-kB p65 and Fos/Jun transcription factors produces potentiated biological function. *EMBO J.* **1 2**:3879-

3891

Stevenson M.S., Haggerty C.A., Lamonica C.M. et al (1990a) Integration is not necessary for expression of human immunodeficiency virus type 1 protein products. *J. Virol.* **6 4**:2421-2425

Stevenson M.S., Stanwick T.L., Dempsey M.P. & Lamonica C.A. (1990b) HIV-1 replication is controlled at the level of T cell activation and proviral integration *EMBO J.* **9**:1551-1560

Stowring L., Haase A.T., Petursson G. *et al.* (1985) Detection of visna virus antigens and RNA in glial cells in foci of demyelination. *Virology* **1 4 1**:311-318

Strebel K., Daugherty D., Clouse K. *et al.* (1987) The HIV 'A' (sor) gene product is essential for virus infectivity. *Nature* **3 2 8**:728-730

Su W., Jackson S., Tijan R. & Echols H. (1991) DNA looping between sites for transcriptional activation: self association of DNA bound Sp1. *Genes & Dev.* **5**:820-826

Suzuki T., Okuno H., Yoshida H. *et al.* (1991) Difference in transcriptional regulatory function between c-Fos and Fra-2. *Nucleic Acids Res.* **1 9**:5537-5542

Suzuki T., Fujisawa J.I., Toita M. & Yoshida M. (1993) The trans activator Tax of human T-cell leukemia virus type I (HTLV-I) interacts with cAMP responsive element (CRE) binding and CRE modulator proteins that bind to the 21 base pair enhancer of HTLV-I. *Proc. Natl. Acad. Sci. USA* **9 0**:610-614

Suzuki T., Hirai H. & Yoshida M. (1994) Tax protein of HTLV-1 interacts with the Rel homology domain of NF-kB p65 and c-REL proteins bound to the theNF-kB binding sites and activates transcription. *Oncogene* **9**:3099-3105

Taddeo B., Haseltine W.A. & Farnet C.M. (1994) Integrase mutants of Human Immunodeficiency Virus type 1 with a specific defect in integration. *J. Virol.* **6 8**:8401-8405

Takada S. & Koike K. (1990) Trans-activation function of a 3' truncated X gene-cell fusion product from integrated Hepatitis B virus DNA in chronic hepatitis tissues. *Proc. Natl. Acad. Sci. USA* **8 7**:5628-5632

Takahashi K., Vigneron M., Matthes H. *et al.* (1986) Requirement for stereospecific alignments for initiation from the simian virus 40 early promoter. *Nature* **3 1 9**:121-126

Tan S-H., Gloss B. & Bernard H-U. (1992) During negative regulation of the human papilomavirus-16 E6 promoter, the viral E2 protein can displace Sp1 from a proximal promter element. *Nuc. Acids Res.* **2 0**:251-256

Thanos D. & Maniatis T. (1992) The high mobility group protein HMGI(Y) is required for NF-kB dependent virus induction of the human IFN-ß gene. *Cell* **7 1**:777-789

Thomas G.H. & Elgin S.C.R. (1988) Protein/DNA architecture of the DNase 1 hypersensitive region of the *Drosophila hsp26* promoter. *EMBO J.* **7**:2191-2201

Thompson F.J., Elder J. & Neil J.C. (1994) Cis- and trans-regulation of feline immunodeficiency virus: identification of functional binding sites in the long terminal repeat. *J. Gen. Virol.* **7 5**:545-554

Tiley L.S. & Cullen B.R. (1992) Structural and functional analysis of the visna virus Rev-response element. *J. Virol.* **6 6**:3609-3615

Tjian R. & Maniatia T. (1994) Transcriptional activation: A complex puzzle with few easy pieces. *Cell* **7 7**:5-8

Tomonaga K., Miyazawa T., Sakuragi J. et al (1993) The feline immunodeficiency virus ORF-A gene facilitates viral replication in established cell lines and peripheral blood lymphocytes. *J. Virol.* **6 7**:5889-5895

Toohey K.L. & Haase A.T. (1994) The rev gene product of visna virus is required for

productive infection. *Virology* **200**:276-280

Travis A. Amsterdam A., Belanger C. & Grosschedl R. (1991) LEF-1, a gene encoding a lymphoid specific protein with an HMG domain, regulates T cell receptor enhancer function. *Genes Dev.* **5**:880-894

Treisman R. (1992) The serum response element. *Trends in Biol. Sci.* **17**:423-426

Tsichlis P.N. & Coffin J.M. (1979) Recombination between the defective component of an acute leukemia virus and rous associated virus 0,an endogenous virus of chickens. *Pro. Natl. Acad. Sci. USA* **76**:3001-3005

Tsukada J., Saito K., Waterman W.R., Webb A.C. & Auron P.E. (1994) Transcription factors NF-IL6 and CREB recognise a common essential site in the human prointerleukin 1ß gene. *Mol. Cell. Biol.* **14**:7285-7297

Twu J.S. & Schloemer (1987) transcriptional trans-activating function of Hepatitis B virus. *J. Virol.* **61**:3448-3453

Twu J.S., Chu K. & Robinson W.S. (1989) Hepatitis B virus X gene activates kB-like enhancer sequences in the long terminal repeat of human immunodeficiency virus I. *Proc. Natl. Acad. Sci. USA* **86**:5168-5172

Twu J.S. & Robinson W.S. (1989) Hepatitis B virus X gene can transactivate heterologous viral sequences. *Proc. Natl. Acad. Sci. USA* **86**:2046-2050

Urban M.B., Schreck R. & Baeuerle P.A. (1991) NF-kB contacts DNA by a heterodimer of the p50 and p65 subunit. *EMBO J.* **10**:1817-1825

Vallet S.M. & Osborne T.F. (1994) Two separate sites contribute to AP-1 activation of the promoter for 3-hydroxy-3-methylglutaryl coenzyme A synthase. *Nuc. Acids Res.* **22**:5184-5189

Van der Molen E.J., Vecht U. & Houwers D.J. (1985) A chronic indurative mastitis in sheep, associated with maedi/visna virus infection. *Vet. Quart.* **7**:112-119

Van Lint C., Burny A. & Verdin E. (1991) The intragenic enhancer of human immunodeficiency virus type 1 contains functional AP-1 binding sites. *J. Virol.* **6 5**:7066-7072

Van Lint C., Ghysdael J., Paras P. Jr., Burny A. & Verdin E. (1994) A transcriptional regulatory element is associated with a nuclease-hypersensitive site in the pol gene of human immunodeficiency virus type-I. *J. Virol.* **6 8**:2632-2648

Vellutini C., Philippon V.,Gambarelli D., *et al.* (1994) The Maedi-Visna virus TAT protein induces multiorgan lymphoid hyperplasia in transgenic mice *J. Virol.* **6 8**:4955-4962

Verdin E., Becker N., Bex F., Droogmans L. & Burny A. (1990) Identification and characterization of an enhancer in the coding region of the genome of human immunodeficiency virus type 1. *Proc. Natl. Acad. Sci. USA* **8 7**:4874-4878

Verdin E. (1991) DNase I-hypersensitive sites are associated with both long terminal repeats and with the intragenic enhancer of integrated human immunodeficiency virus type 1. *J. Virol.* **6 5**:6790-6799

Verdin E., Paras P. Jr., Van Lint C. (1993) Chromatin disruption in the promoter of human immunodeficiency virus type 1 during transcriptional activation. *EMBO J.* **1 2**:3249-3259

Vigne R., Filippi P., Querat G. *et al.* (1985) Molecular biology of maedi/visna viruses.pp27-44 *in* Slow Viruses in Sheep, Goats and Cattle. Commission of the European Communities.

Vigne R., Barban V., Querat G. *et al.* (1987) Transcription of visna virus durin its lytic cycle: evidence for a sequential early and late gene expression. *Virology* **1 6 1**:218-227

Vigne R., Gourdou I., Mazarin V. *et al.* (1990) Regulatory genes of visna virus. *Develop. Biol.*

Standard. **7 2**:213-222

Wagner S. & Green M. (1993) HTLV-I Tax protein stimulation of DNA binding bZIP ptoteins by enhancing dimerisation. *Science* **2 6 2**:395-399

Wang E.H. & Tijan R. (1994) Promoter selective transcriptional defect in cell cycle mutant ts13 rescued by hTAFII250. *Science* **2 6 3**:811-814

Waterman M.L. & Jones K.A. (1990) Purification of TCR-1α, a T-cell-specific transcription factor that activates the T-cell receptor Cα gene enhancer in a context dependent manner. *New Biol.* **2**:621-636

Wefald F.C., Devlin B.H., & Williams R.S. (1990) Functional heterogeneity of mammalian TATA box sequences revealed by interaction with a cell specific enhancer. *Nature* **3 4 4**:260-262

Westervelt P., Trowbridge D.B. Epstein L.G. *et al.* (1992) Macrophage tropism determinants of human immunodeficiency virus type 1 in vivo. *J. Virol.* **6 6**:2577-2582

Wilkinson D.G., Bhatt S., Ryseck R-P. & Bravo R. (1989) Tissue-specific of c-jun and junB during organogenesis in the mouse. *Development* **1 0 6**:465-471

Winslow B.J., Pomerantz R.J., Bagasra O. & Trono D. (1993) HIV-1 latency due to the site of proviral integration. *Virology* **1 9 6**:849-854

Wiskerchen M. & Muesing M.A. (1995) Human immunodeficiency virus type I integrase: effects of mutations on viral ability to integrate, direct viral gene expression from unintegrated viral DNA templates, and sustain viral propagation in primary cells. *J. Virol.* **6 9**:376-386

Xanthoudakis S., Miao G., Wang F., Pan Y-C E. & Curran T (1992) Redox ativation of Fos-Jun DNA binding activity is mediated by a DNA repair enzyme. *EMBO J.* **11**:3323-3335

Yang- Yen H-F., Chambard J-C., Sun Y-L. *et al.* (1990) Transcriptional interference between c-

Jun and the glucocorticoid receptor: mutual inhibition of DNA binding due to direct protein-protein interaction. *Cell* **6 2**:1205-1215

Yoshida T., Suzuki T, Sato H., Nishina H. & Iba H. (1993) Analysis of Fra-2 gene expression. *Nucleic Acids Res.* **2 1**:2715-2721

Zhao D.X., Taraboulos A., Ou J.H. & Yen T.S.B. (1990) Activation of class I major histocompatibility complex gene expression by Hepatitis B virus. *J. Virol.* **6 4**:4025-4028

Zhao L.J. & Giam C.Z. (1991) Interaction of the human T-cell lymphotropic virus type-I (HTLV-I) transcriptional activator Tax with cellular factors that bind specifically to the 21 base pair repeats in the HTLV-I enhancer. *Proc. Natl. Acad. Sci. USA* **8 8**:11445-11449

Zhao L.J. & Giam C.Z. (1992) Human T-cell lymphotropic virus type-I (HTLV-I) transcriptional activator, Tax, enhances CREB binding to HTLY-I 21 base-pair repeats by protein-protein interaction. *Proc. Natl. Acad. Sci. USA* **8 9**:7070-7074